

Differential expression of key CmPn members distinguishes
histological and immune subtypes of hepatic cancers

Johnathan Abou-Fadel, Victoria Reid, Alexander Le, Jun Zhang *

Department of Molecular and Translational Medicine (MTM)

Texas Tech University Health Science Center El Paso, El Paso, TX 79905 USA

*Correspondence: jun.zhang2000@gmail.com; Tel.: +1-(915)-215-4197

Legends

Suppl. Figures

Suppl. Figure S1. Differentially expressed genes across immune subtypes for Hepatocellular Carcinoma (HCC). **A)** *Comparisons between wound healing (C1) and IFN- γ dominant (C2) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of differentially expressed genes (DEGs). Results of the gene ontology enrichment analysis comparing C1 and C2 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **B)** *Comparisons between wound healing (C1) and inflammatory (C3) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C1 and C3 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **C)** *Comparisons between wound healing (C1) and lymphocyte depleted (C4) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C1 and C4 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **D)** *Comparisons between IFN- γ dominant (C2) and inflammatory (C3) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C2 and C3 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **E)** *Comparisons between IFN- γ dominant (C2) and lymphocyte depleted (C4) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C2 and C4 are provided for biological processes (lower left panel), molecular functions

(upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **F)** *Comparisons between inflammatory (C3) and lymphocyte depleted (C4) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C3 and C4 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). For all panels, scatter plots display the log₂-fold changes and statistical significance of each gene calculated by performing a differential gene expression analysis. Genes in volcano plots with log₂FC>1.5 and p-value<0.05 are in red while genes with log₂FC<-1.5 and p-value<0.05 are in blue. All enrichment analysis results were generated using Enrichr where the x axis indicates the -log₁₀(P-value); the top 15 pathways are provided and color-coded (red-up/blue-down) and pathways with p-values<0.05 and FDR<0.1 are bolded.

Suppl. Figure S2. Differentially expressed genes across immune subtypes for Cholangiocarcinoma (CCA). **A)** *Comparisons between wound healing (C1) and IFN- γ dominant (C2) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of differentially expressed genes (DEGs). Results of the gene ontology enrichment analysis comparing C1 and C2 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **B)** *Comparisons between wound healing (C1) and inflammatory (C3) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C1 and C3 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **C)** *Comparisons between wound healing (C1) and lymphocyte depleted (C4) subtypes.* Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of

the gene ontology enrichment analysis comparing C1 and C4 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **D) Comparisons between *IFN- γ* dominant (C2) and inflammatory (C3) subtypes.** Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C2 and C3 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **E) Comparisons between *IFN- γ* dominant (C2) and lymphocyte depleted (C4) subtypes.** Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C2 and C4 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). **F) Comparisons between inflammatory (C3) and lymphocyte depleted (C4) subtypes.** Sample size is shown for each comparing group (top left panel) along with a scatter plot (middle left panel) of DEGs. Results of the gene ontology enrichment analysis comparing C3 and C4 are provided for biological processes (lower left panel), molecular functions (upper right panel), reactome pathways (middle right panel), and KEGG pathways (lower right panel). For all panels, scatter plots display the log₂-fold changes and statistical significance of each gene calculated by performing a differential gene expression analysis. Genes in volcano plots with log₂FC>1.5 and p-value<0.05 are in red while genes with log₂FC<-1.5 and p-value<0.05 are in blue. All enrichment analysis results were generated using Enrichr where the x axis indicates the -log₁₀(P-value); the top 15 pathways are provided and color-coded (red-up/blue-down) and pathways with p-values<0.05 and FDR<0.1 are bolded.

Suppl. Figure S3. RNAseq expression profiling for key CmPn players integrating sociological, follow-up, and clinical survival data for hepatic cancers: We investigated key CSC, mPRs and nPRs (CmPn) players expression analysis (along with established liver cancer biomarker *AFP*) using demographic and follow-up

data to assess differential expression across major races and evaluate impact on tumor recurrence following treatment. **A)** Differential expression for key CmPn members, along with *AFP*, among the three most prevalent races diagnosed with CCAs (n=45). **B)** Differential expression for key CmPn members, along with *AFP*, among the five major locations for tumor recurrence in patients diagnosed with CCAs (n=20). **C)** Differential expression for key CmPn members, along with *AFP*, based on vital status for patients diagnosed with CCAs (n=45). **D1-2)** Differential expression of key CmPn members, along with *AFP*, for patients diagnosed with HCC (**D-1**, n=412) or patients diagnosed with CCAs (**D-2**, n=45). **E1-2)** Differential expression for key CmPn members, along with *AFP*, for residual tumor classification in patients diagnosed with HCCs (**E-1**, n=380) or for patients diagnosed with CCAs (**E-2**, n=41). For all graphs, X axis details genes profiled, while Y axis details Log2 batch effects normalized RNAseq expression data. All graphs were produced using the Xena platform.

Suppl. Figure S4. Hematoxylin and Eosin (H&E) imaging of microarray slides for normal and seven subtypes of hepatic cancers. H&E pre-stained microarray slides were purchased to visualize tissue patterns and subcellular structural differences between lesion subtypes. A representative normal healthy liver tissue (NORM) is displayed (Panel A) along with a representative hepatocellular carcinoma (HCC) tissue (panel B), cholangiocarcinoma (CCA) tissue (panel C), combined hepatocholangiocarcinoma (cHCC-CCA) tissue (panel D), adenosquamous carcinoma (ASC) tissue (panel E), undifferentiated pleomorphic sarcoma (UPS) tissue (panel F), clear cell carcinoma (CCC) tissue (panel G) and hepatic angiosarcoma (HAS) tissue (panel H). Tissues were imaged using a Nikon EclipseTi microscope with Images acquired using a 20X objective lens to ensure coverage of tissue.

Suppl. Figure S5. Immunofluorescence imaging of AFP and nPRs for hepatic cancers. Comparative protein expression patterns for AFP and nPRs were measured with immunofluorescence-labeled antibodies,

normalized against nuclear staining (DAPI), and quantified using Nikon Elements Analysis software. **A)** Representative IF images of AFP protein expression between NORM and CCA tissues is displayed, along with normalized quantification for all images. **B)** Normalized quantification for all images of nPRs protein expression among different liver cancer types and NORM tissue (images not shown).

Suppl. Tables

Suppl. Table S1. Databases used for NCI-GDC clinical profiling of hepatic cancers in Figure 1. Databases used by the Genomic Data Commons (GDC) data portal from the National Cancer Institute (NCI), are detailed in tab 1 for the general analysis of hepatic cancers without any gene filters (general clinical observations) or tab 2 for the filtered analysis to only analyze patient samples with differential expression for any of the key CmPn players (CmPn network associated clinical observations). The database of Genotypes and Phenotypes (dbGaP) study accession number is provided for each database.

Suppl. Table S2A. Raw RNAseq/enrichment results for normal healthy patients compared to HCC patients. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Figure 2C-1 is provided in tab 1 with gene symbol in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding KEGG pathways enrichment analysis performed using Enrichr is provided in tab 4. For Tabs 2-4, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up-regulated), respectively, or columns I-N (down-regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are

color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S2B. Raw RNAseq/enrichment results for normal healthy patients compared to CCA patients. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Figure 2D-1 is provided in tab 1 with gene symbol in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding KEGG pathways enrichment analysis performed using Enrichr is provided in tab 4. For Tabs 2-4, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up-regulated), respectively, or columns I-N (down-regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S2C. Raw RNAseq/enrichment results for HCC compared to CCA patients. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Figure 2E-1 is provided in tab 1 with gene symbol in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding KEGG pathways enrichment analysis performed using Enrichr is provided in tab 4. For Tabs 2-4, the terms, p-value, FDR, Z-score, combined score and corresponding genes are

provided in columns A-F (up-regulated), respectively, or columns I-N (down-regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S3. Antibodies and Microarray slides used for Immunofluorescence / H&E analysis.

Details regarding application, antigen, gene, clone, manufacturer, catalog number, and primary and secondary dilutions (if applicable) for primary and secondary antibodies used are listed in tab 1. Information about microarray slides purchased for imaging are detailed in tab 2 with catalog #, details, manufacturer, and application. Abbreviations: IF, immunofluorescence; IHC, immunohistochemistry.

Suppl. Table S4A. Raw RNAseq/enrichment results for HCC Immune subtypes C1 v C2. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 1A is provided in tab 1; gene symbol is in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-values in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up-regulated), respectively, or columns I-N (down-regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S4B. Raw RNAseq/enrichment results for HCC Immune subtypes C1 v C3. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 1B is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S4C. Raw RNAseq/enrichment results for HCC Immune subtypes C1 v C4. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 1C is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which

are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S4D. Raw RNAseq/enrichment results for HCC Immune subtypes C2 v C3. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 1D is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S4E. Raw RNAseq/enrichment results for HCC Immune subtypes C2 v C4. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 1E is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs

2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S4F. Raw RNAseq/enrichment results for HCC Immune subtypes C3 v C4. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 1F is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S5A. Raw RNAseq/enrichment results for CCA Immune subtypes C1 v C2. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 2A is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-values in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular

functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up-regulated), respectively, or columns I-N (down-regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and $FDR < 0.1$ are bolded.

Suppl. Table S5B. Raw RNAseq/enrichment results for CCA Immune subtypes C1 v C3. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 2B is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and $FDR < 0.1$ are bolded.

Suppl. Table S5C. Raw RNAseq/enrichment results for CCA Immune subtypes C1 v C4. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 2C is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column

E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S5D. Raw RNAseq/enrichment results for CCA Immune subtypes C2 v C3. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 2D is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

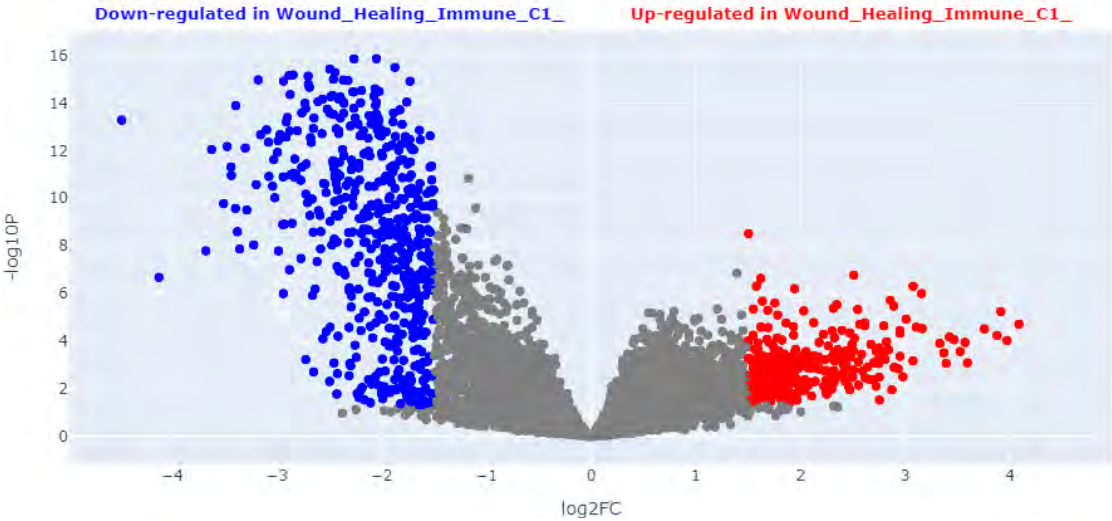
Suppl. Table S5E. Raw RNAseq/enrichment results for CCA Immune subtypes C2 v C4. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with

CmPn expression differences) in Suppl. Fig 2E is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

Suppl. Table S5F. Raw RNAseq/enrichment results for CCA Immune subtypes C3 v C4. The raw RNAseq data used to generate the differential expression analysis (volcano plot and corresponding table with CmPn expression differences) in Suppl. Fig 2F is provided in tab 1; gene symbol is located in column A, log fold changes (logFC) in column B, average expression in column C, t-values in column D, P-value in column E and adjusted P-values in column F. The raw data for the corresponding Gene Ontology enrichment analysis performed using Enrichr is provided in tab 2 (Biological processes) and tab 3 (Molecular functions). The raw data for the corresponding signaling pathways enrichment analysis performed using Enrichr is provided in tab 4 (reactome signaling pathways) and tab 5 (KEGG signaling pathways). For tabs 2-5, the terms, p-value, FDR, Z-score, combined score and corresponding genes are provided in columns A-F (up regulated), respectively, or columns I-N (down regulated), respectively. The top 15 terms (which are illustrated in the corresponding figure) are color-coded (red for up regulated or blue for down regulated) and terms that had p-values <0.05 and FDR<0.1 are bolded.

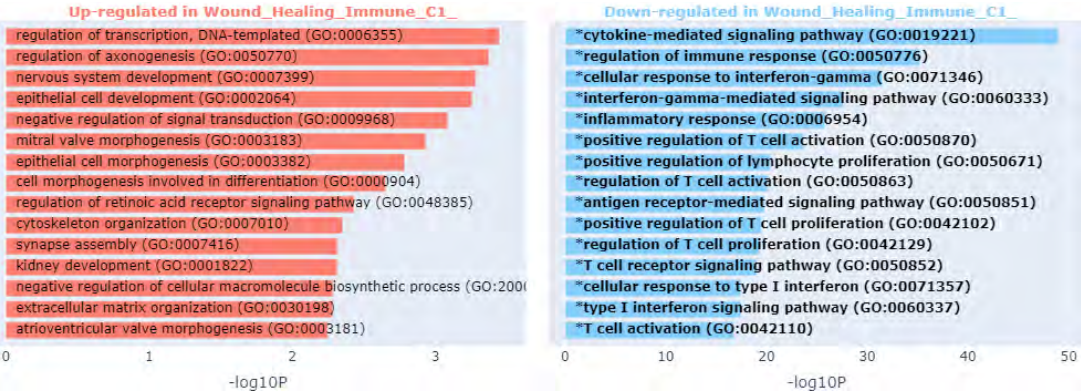
S1A

Immunological Subtypes	Sample size
Wound Healing (C1)	22
IFN-γ dominant (C2)	45

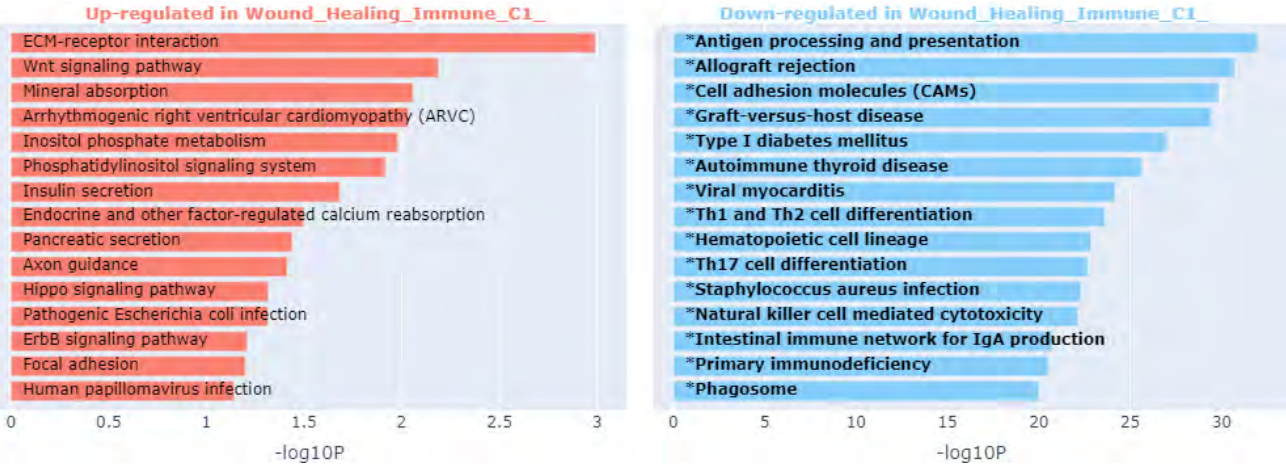


Gene symbol	logFC	AveExpr	t	PValue
<i>CCM1</i>	0.518192	3.930383	3.14197	2.00E-03
<i>PAQR6</i>	0.68546	1.056312	2.055773	4.00E-02

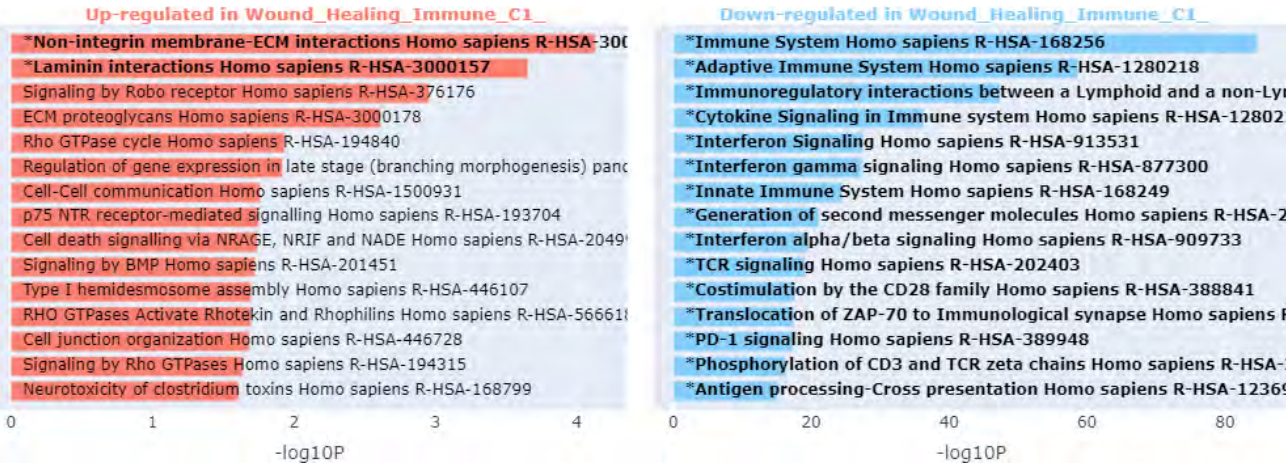
Biological Processes



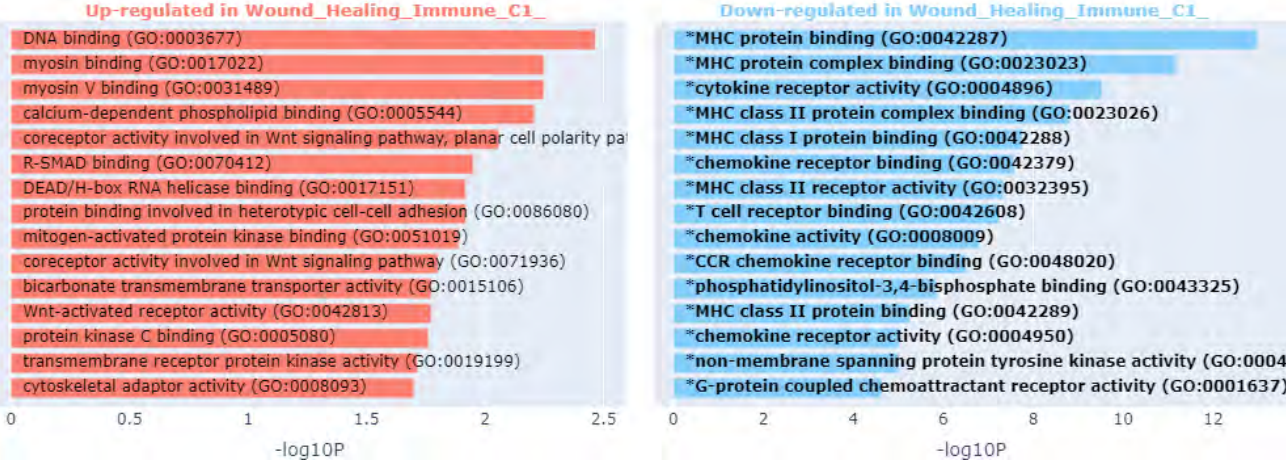
KEGG Pathways



Reactome pathways

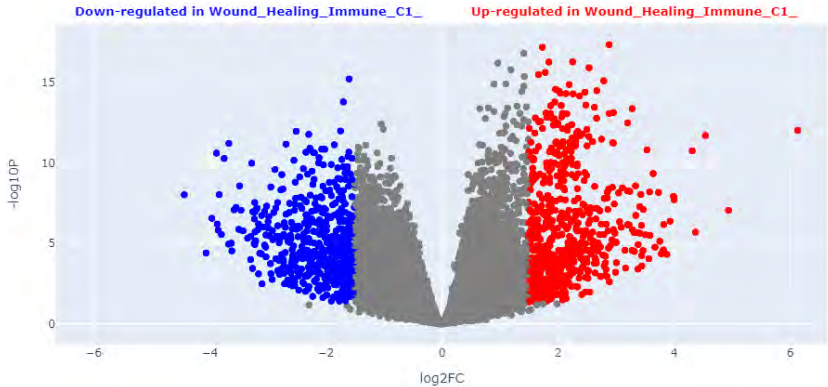


Molecular Functions



S1B

Immunological Subtypes	Sample size
Wound Healing (C1)	22
Inflammatory (C3)	135



Gene symbol	logFC	AveExpr	t	P.Value
<i>PAQR6</i>	1.191942	0.344273	3.789672	2.14E-04
<i>AFP</i>	2.896023	3.312498	3.089215	2.37E-03
<i>PAQR5</i>	1.484179	2.100832	2.540652	1.20E-02
<i>CCM3</i>	0.200158	3.992807	2.209856	2.85E-02
<i>PGRMC1</i>	-1.14256	8.277952	-5.10998	9.15E-07
<i>PGRMC2</i>	-0.49476	6.458132	-3.84915	1.71E-04

Biological Processes

Up-regulated in Wound_Healing_Immune_C1

- *DNA metabolic process (GO:0006259)
- *DNA replication (GO:0006260)
- *mitotic cell cycle phase transition (GO:0044772)
- *mitotic sister chromatid segregation (GO:0000070)
- *mitotic spindle organization (GO:0007052)
- *cellular macromolecule biosynthetic process (GO:0034)
- *microtubule cytoskeleton organization involved in mitc
- *centromere complex assembly (GO:0034508)
- *DNA repair (GO:0006281)
- *G1/S transition of mitotic cell cycle (GO:0000082)
- *chromatin remodeling at centromere (GO:0031055)
- *cell cycle G1/S phase transition (GO:0044843)
- *CENP-A containing nucleosome assembly (GO:0034080)
- *CENP-A containing chromatin organization (GO:006164)
- *cellular response to DNA damage stimulus (GO:000697)

Down-regulated in Wound_Healing_Immune_C1

- *fatty acid metabolic process (GO:0006631)
- *alpha-amino acid catabolic process (GO:1901606)
- *fatty acid catabolic process (GO:0009062)
- *branched-chain amino acid metabolic process (GO:000)
- *monocarboxylic acid metabolic process (GO:0032787)
- *cellular amino acid catabolic process (GO:0009063)
- *dicarboxylic acid metabolic process (GO:0043648)
- *fatty acid oxidation (GO:0019395)
- *acyl-CoA metabolic process (GO:0006637)
- *lipid transport (GO:0006869)
- *coenzyme metabolic process (GO:0006732)
- *branched-chain amino acid catabolic process (GO:0000)
- *short-chain fatty acid catabolic process (GO:0019626)
- *sulfur amino acid metabolic process (GO:0000096)
- *short-chain fatty acid metabolic process (GO:0046459)

KEGG Pathways

Up-regulated in Wound_Healing_Immune_C1

- *Cell cycle
- *DNA replication
- *Oocyte meiosis
- *Homologous recombination
- *Progesterone-mediated oocyte maturation
- *Spliceosome
- *Fanconi anemia pathway
- *Pyrimidine metabolism
- *Mismatch repair
- *Human T-cell leukemia virus 1 infection
- *Cellular senescence
- *p53 signaling pathway
- *Small cell lung cancer
- *Pathogenic Escherichia coli infection
- *RNA transport

Down-regulated in Wound_Healing_Immune_C1

- *Complement and coagulation cascades
- *Fatty acid degradation
- *Valine, leucine and isoleucine degradation
- *Tyrosine metabolism
- *Peroxisome
- *Propanoate metabolism
- *Butanoate metabolism
- *Glyoxylate and dicarboxylate metabolism
- *beta-Alanine metabolism
- *Tryptophan metabolism
- *Glycolysis / Gluconeogenesis
- *PPAR signaling pathway
- *Pyruvate metabolism
- *Glycine, serine and threonine metabolism
- *Drug metabolism

Molecular Functions

Up-regulated in Wound_Healing_Immune_C1

- *DNA binding (GO:0003677)
- *DNA-dependent ATPase activity (GO:0008094)
- *tubulin binding (GO:0015631)
- *DNA helicase activity (GO:0003678)
- *microtubule binding (GO:0008017)
- *single-stranded DNA-dependent ATPase activity (GO:0003777)
- *microtubule motor activity (GO:0003777)
- *protein-DNA loading ATPase activity (GO:0033170)
- *DNA clamp loader activity (GO:0003689)
- *ATP-dependent DNA helicase activity (GO:0004003)
- *3'-5' DNA helicase activity (GO:0043138)
- *RNA binding (GO:0003723)
- *double-stranded DNA binding (GO:0003690)
- *motor activity (GO:0003774)
- *DNA secondary structure binding (GO:0000217)

Down-regulated in Wound_Healing_Immune_C1

- *oxidoreductase activity, acting on the CH-OH group of
- *alcohol dehydrogenase (NAD) activity (GO:0004022)
- *alcohol dehydrogenase activity, zinc-dependent (GO:0004022)
- *oxidoreductase activity, acting on the aldehyde or oxo
- *oxidoreductase activity, acting on paired donors, with
- *NAD binding (GO:0051287)
- *acyl-CoA dehydrogenase activity (GO:0003995)
- *fatty acid ligase activity (GO:0015645)
- *zinc ion binding (GO:0008270)
- *oxidoreductase activity, acting on single donors with i
- *C-acyltransferase activity (GO:0016408)
- *transition metal ion binding (GO:0046914)
- *aldehyde dehydrogenase (NAD) activity (GO:0004029)
- *fatty-acyl-CoA synthase activity (GO:0004321)
- *acyl-CoA ligase activity (GO:0003996)

Reactome pathways

Up-regulated in Wound_Healing_Immune_C1

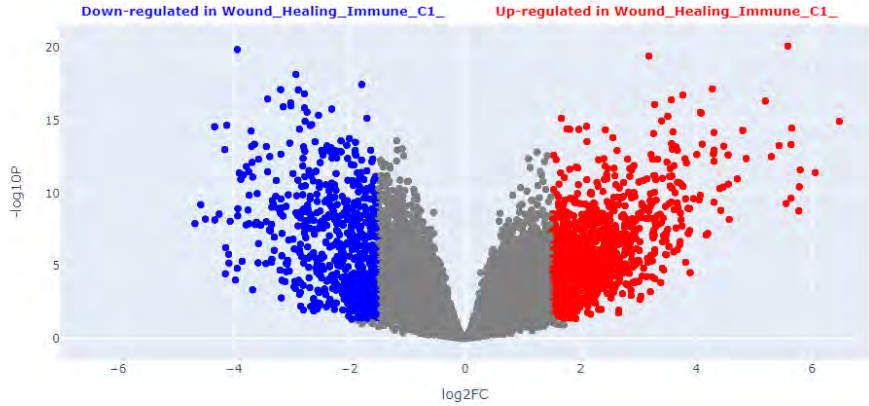
- *Cell Cycle Homo sapiens R-HSA-1640170
- *Cell Cycle, Mitotic Homo sapiens R-HSA-69278
- *M Phase Homo sapiens R-HSA-68886
- *Mitotic Prometaphase Homo sapiens R-HSA-68877
- *Resolution of Sister Chromatid Cohesion Homo sapiens
- *Mitotic Metaphase and Anaphase Homo sapiens R-HSA-
- *RHO GTPases Activate Formins Homo sapiens R-HSA-5
- *Mitotic Anaphase Homo sapiens R-HSA-68882
- *Separation of Sister Chromatids Homo sapiens R-HSA-
- *Cell Cycle Checkpoints Homo sapiens R-HSA-69620
- *RHO GTPase Effectors Homo sapiens R-HSA-195258
- *Signaling by Rho GTPases Homo sapiens R-HSA-19431
- *Activation of ATR in response to replication stress Hom
- *Mitotic G1-G1/S phases Homo sapiens R-HSA-453279
- *G2/M Checkpoints Homo sapiens R-HSA-69481

Down-regulated in Wound_Healing_Immune_C1

- *Metabolism Homo sapiens R-HSA-1430728
- *Metabolism of lipids and lipoproteins Homo sapiens R-
- *Biological oxidations Homo sapiens R-HSA-211859
- *Phase 1 - Functionalization of compounds Homo sapier
- *Metabolism of amino acids and derivatives Homo sapie
- *Branched-chain amino acid catabolism Homo sapiens R-
- *Intrinsic Pathway of Fibrin Clot Formation Homo sapie
- *Regulation of Complement cascade Homo sapiens R-HS
- *Ethanol oxidation Homo sapiens R-HSA-71384
- *Formation of Fibrin Clot (Clotting Cascade) Homo sapie
- *Fatty acid, triacylglycerol, and ketone body metabolism
- *Histidine, lysine, phenylalanine, tyrosine, proline and t
- *Complement cascade Homo sapiens R-HSA-166658
- *Metabolism of vitamins and cofactors Homo sapiens R-
- *Metabolism of water-soluble vitamins and cofactors Ho

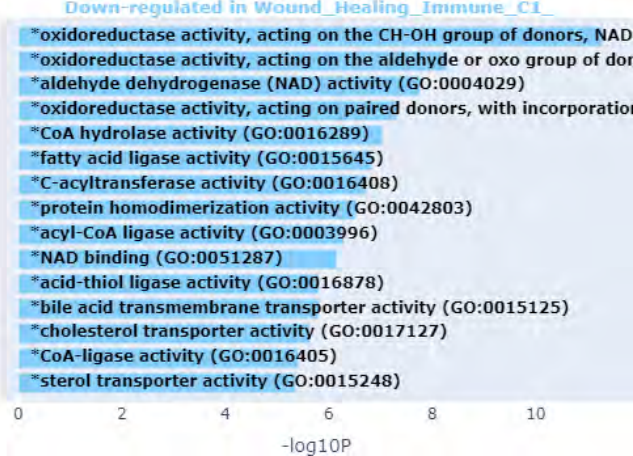
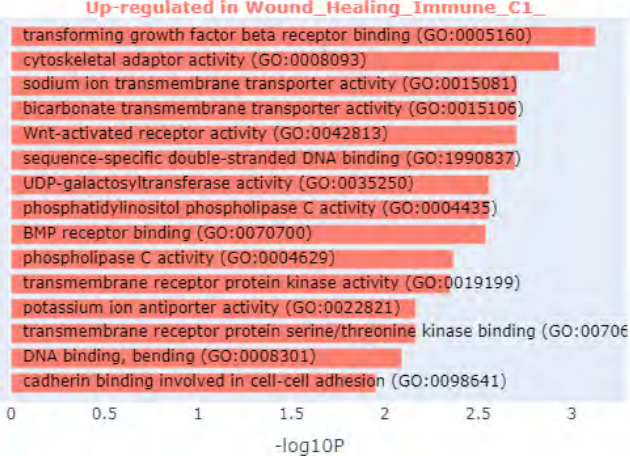
S1C

Immunological Subtypes	Sample size
Wound Healing (C1)	22
Lymphocyte depleted (C4)	159

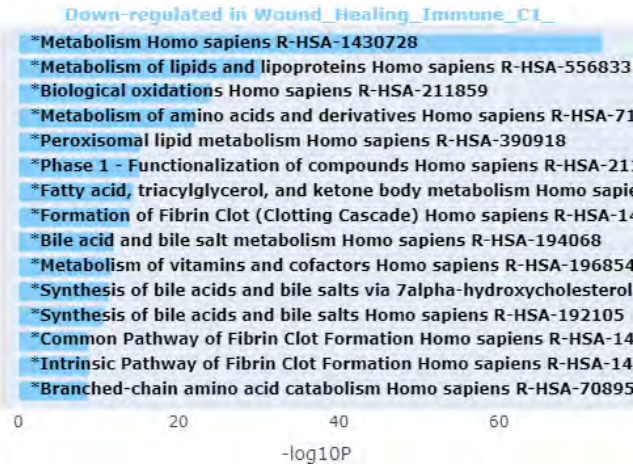
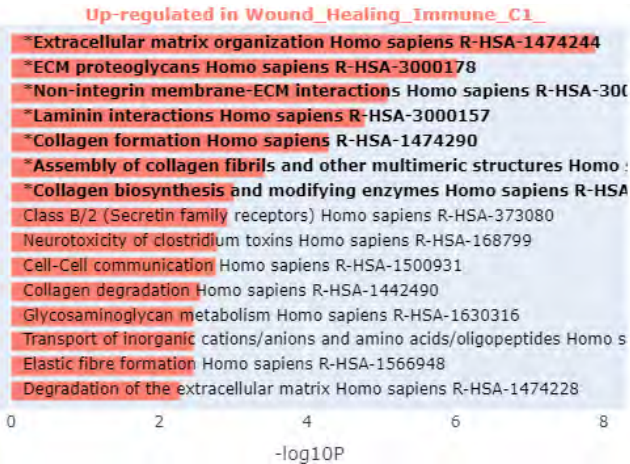


Gene symbol	logFC	AveExpr	t	P.Value
AFP	2.882463	3.183769	3.563241	4.67E-04
PAQR8	1.154065	1.557629	3.544084	5.00E-04
PAQR5	2.228739	1.326255	3.43325	7.37E-04
PAQR6	1.054751	0.347308	3.42513	7.58E-04
PGRMC1	-1.33634	8.378224	-6.01506	9.56E-09
PGRMC2	-0.64401	6.503715	-5.01827	1.23E-06
PAQR9	-0.8919	4.253548	-3.13713	1.99E-03
CCM2	-0.3834	6.047957	-2.40294	1.73E-02

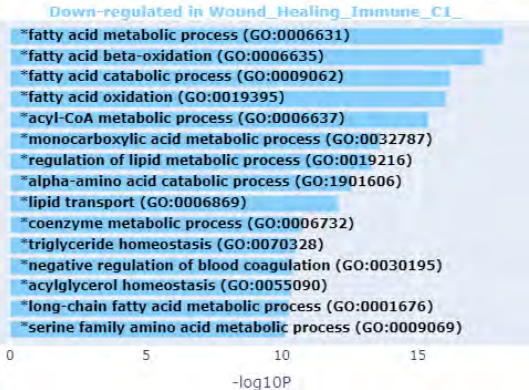
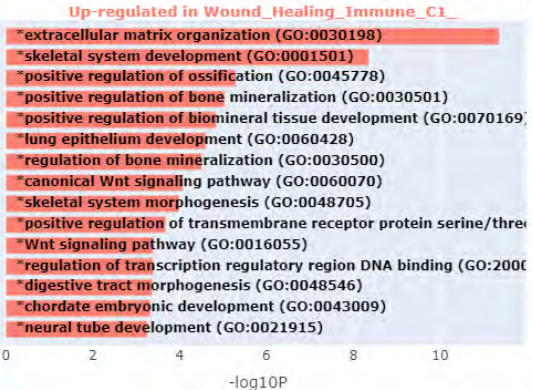
Molecular Functions



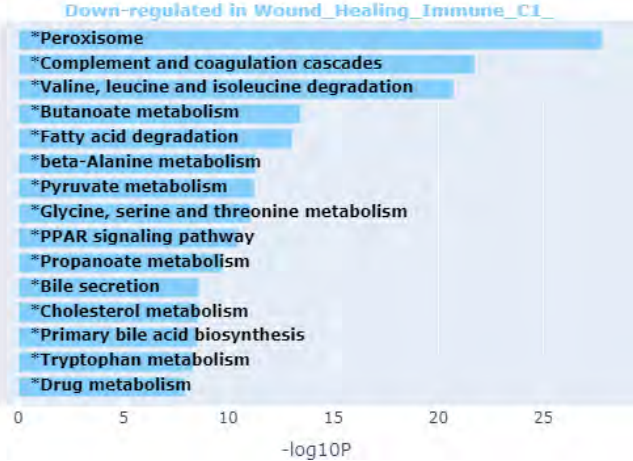
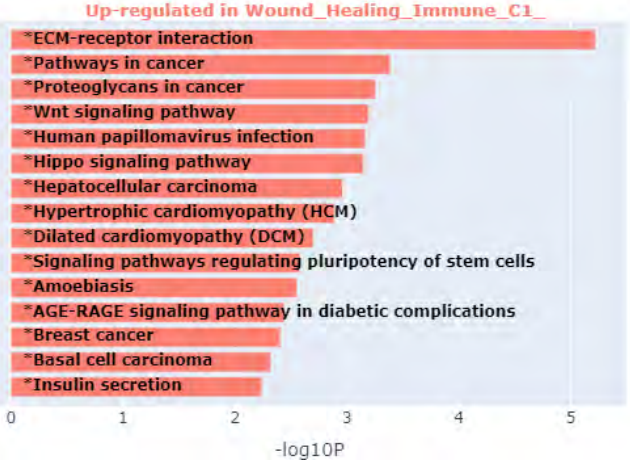
Reactome pathways



Biological Processes

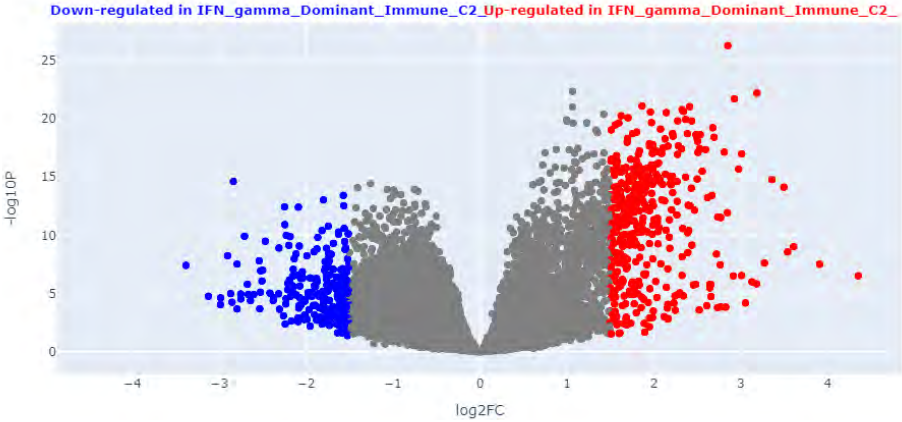


KEGG Pathways



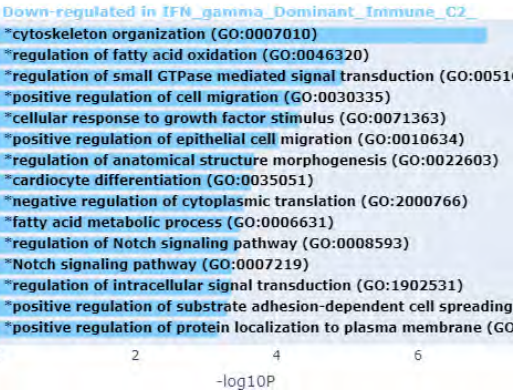
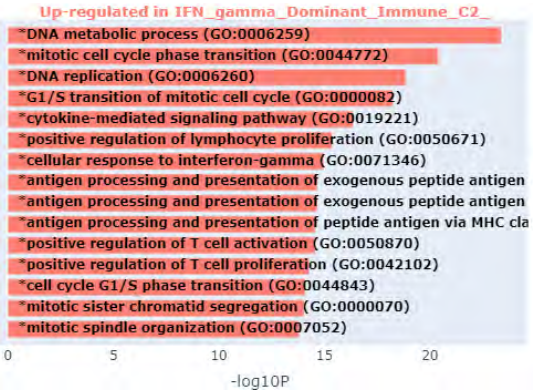
S1D

Immunological Subtypes	Sample size
IFN-γ dominant (C2)	45
Inflammatory (C3)	135

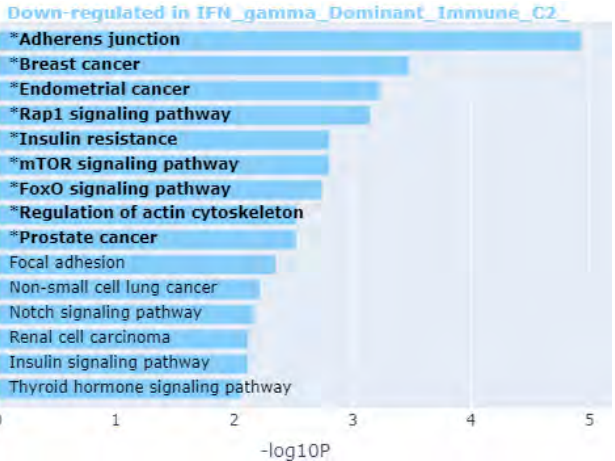
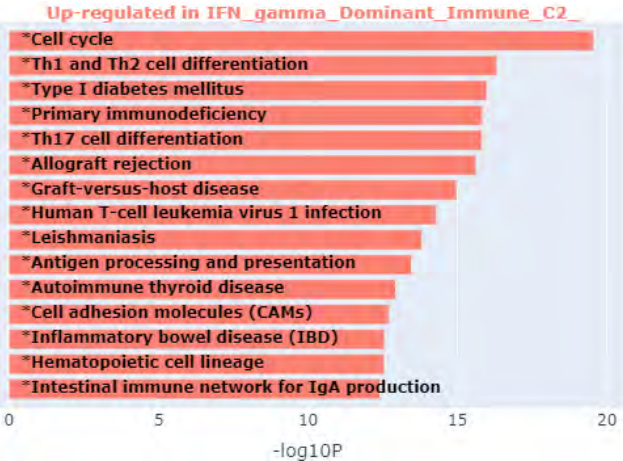


Gene symbol	logFC	AveExpr	t	P.Value
AFP	3.916882	3.872694	5.782277	3.16E-08
CCM3	0.153459	3.986799	2.163984	3.18E-02
PAQR6	0.500112	0.287465	2.111144	3.61E-02
PAQR8	0.42792	2.253244	2.005111	4.64E-02
PGRMC1	-0.7814	8.228141	-4.66774	5.90E-06
CCM1	-0.47107	3.94761	-4.57818	8.68E-06
nPRs	-0.98985	-1.40387	-2.97771	3.30E-03
PGRMC2	-0.29999	6.436494	-2.93692	3.74E-03

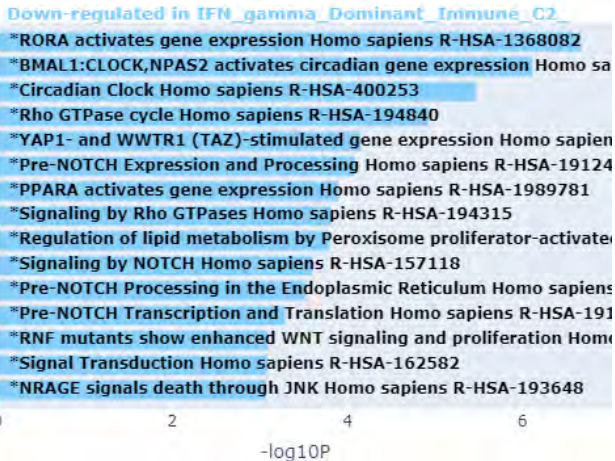
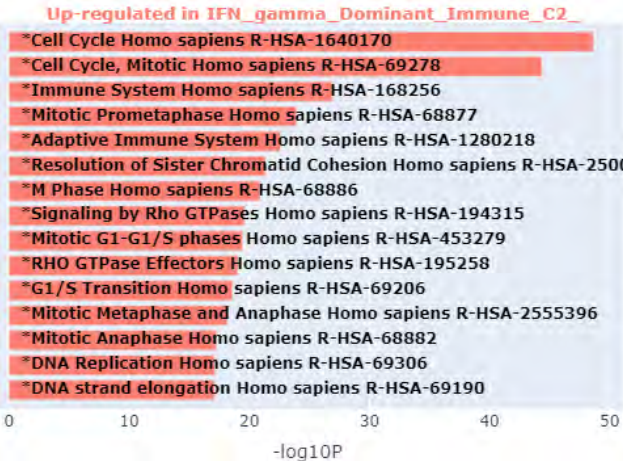
Biological Processes



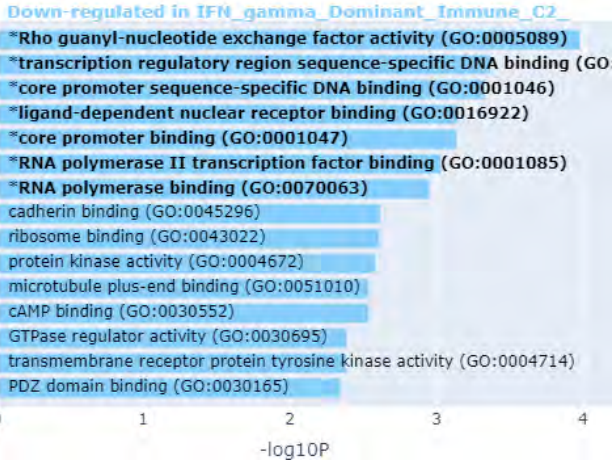
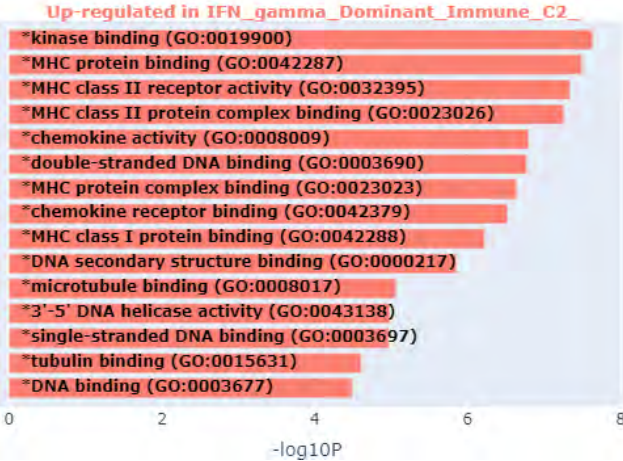
KEGG Pathways



Reactome pathways

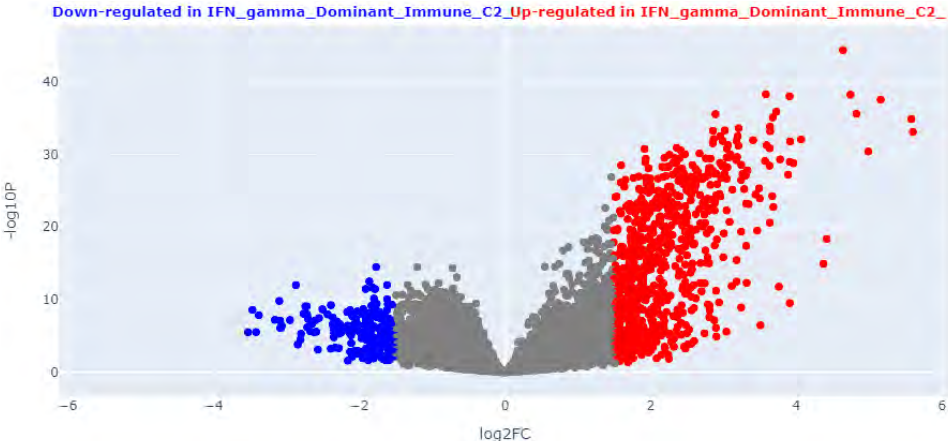


Molecular Functions



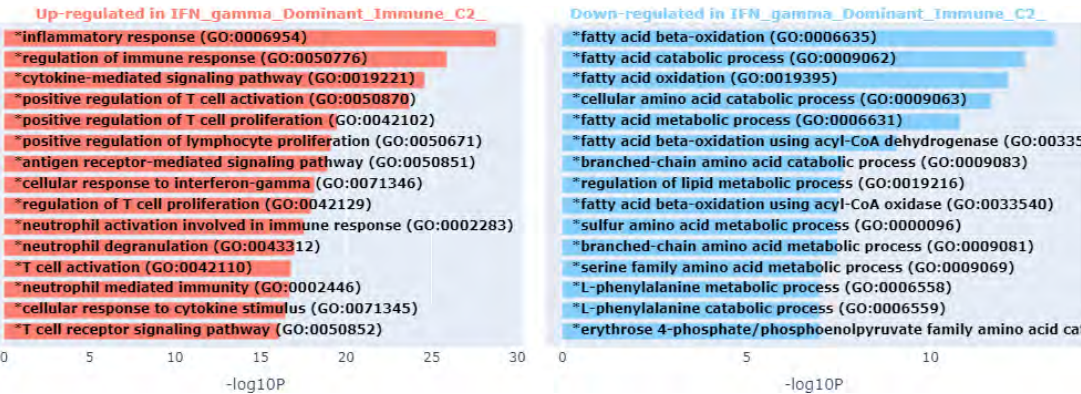
S1E

Immunological Subtypes	Sample size
IFN-γ dominant (C2)	45
Lymphocyte depleted (C4)	159

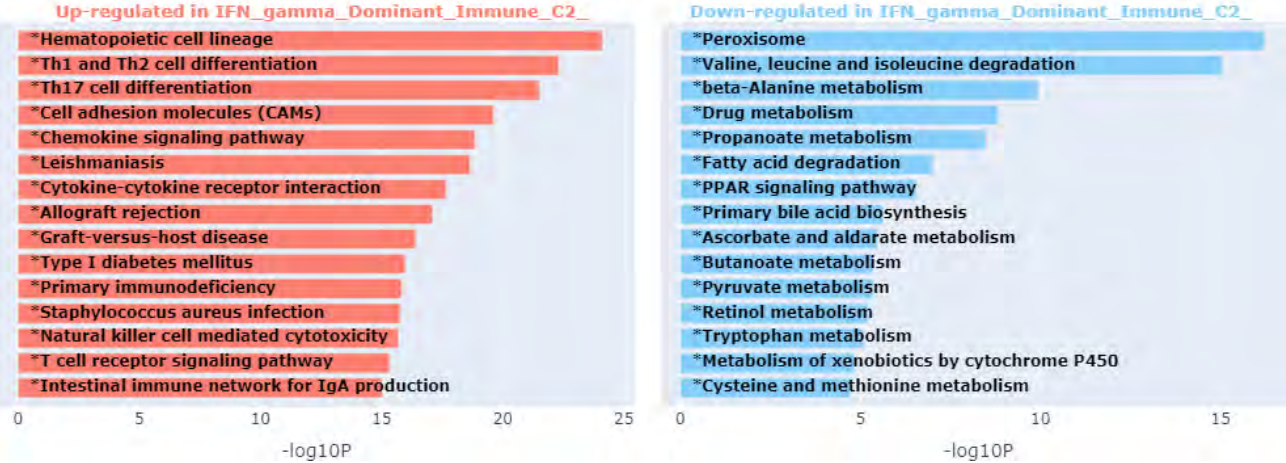


Gene symbol	logFC	AveExpr	t	P.Value
<i>AFP</i>	3.921464	3.692572	6.630822	2.88E-10
<i>PAQR8</i>	1.116469	1.652463	4.713229	4.48E-06
<i>PAQR5</i>	1.129452	1.295504	2.361378	1.91E-02
<i>CCM1</i>	-0.37457	3.833406	-3.77653	2.08E-04
<i>PGRMC1</i>	-0.9364	8.322968	-5.7336	3.47E-08
<i>PGRMC2</i>	-0.40891	6.479483	-4.12465	5.39E-05

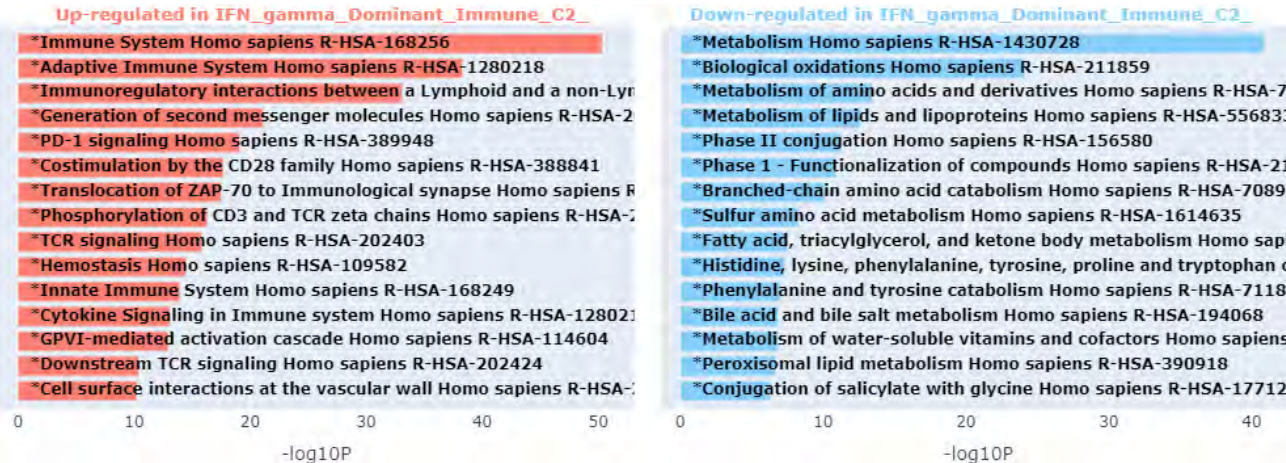
Biological Processes



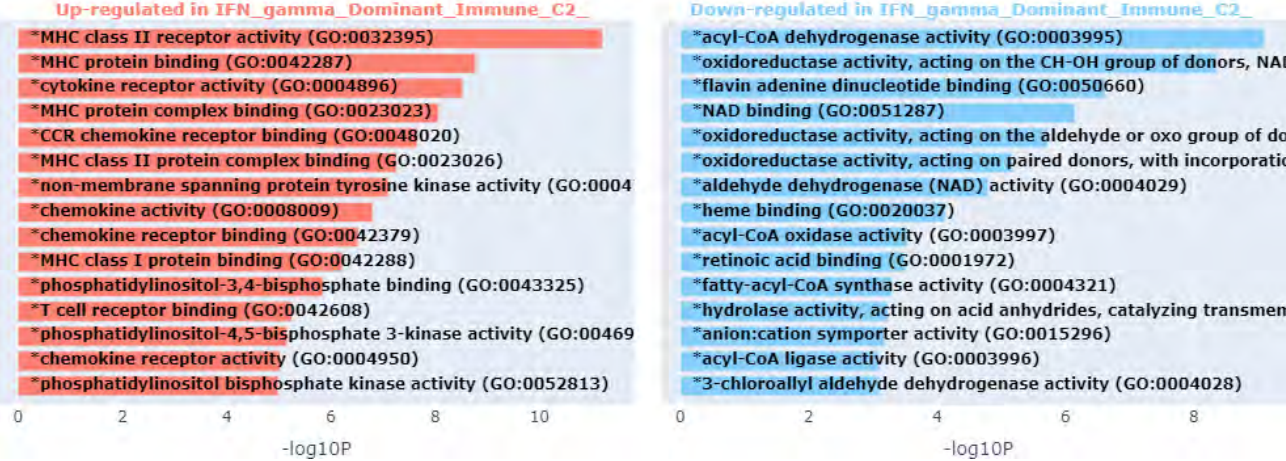
KEGG Pathways



Reactome pathways

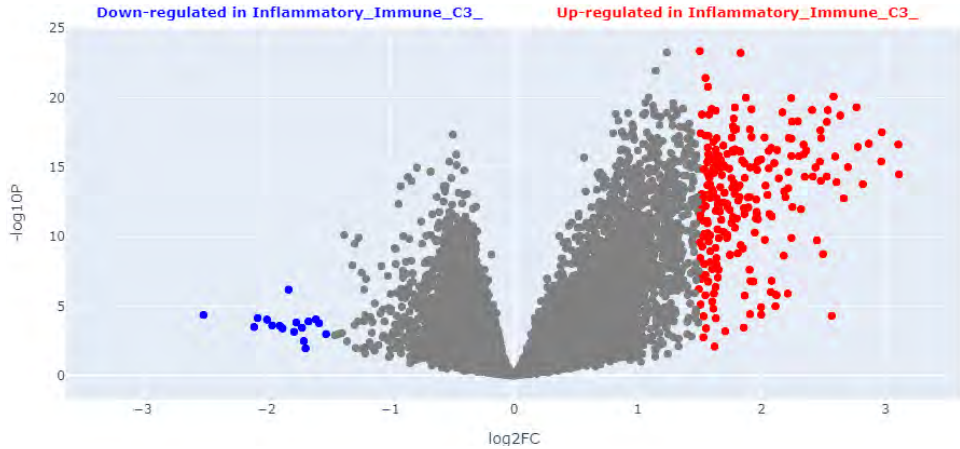


Molecular Functions



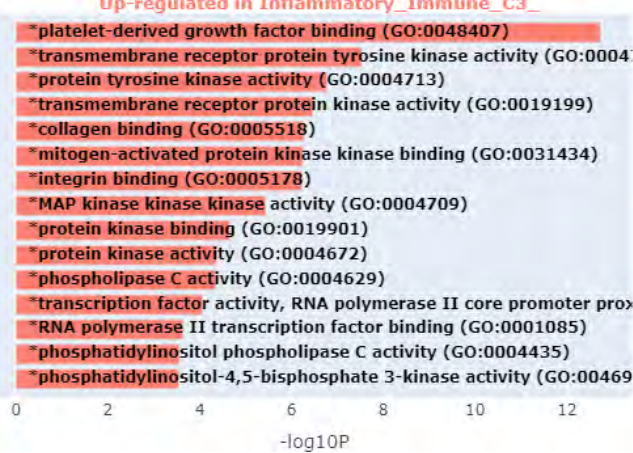
S1F

Immunological Subtypes	Sample size
Inflammatory (C3)	135
Lymphocyte depleted (C4)	159

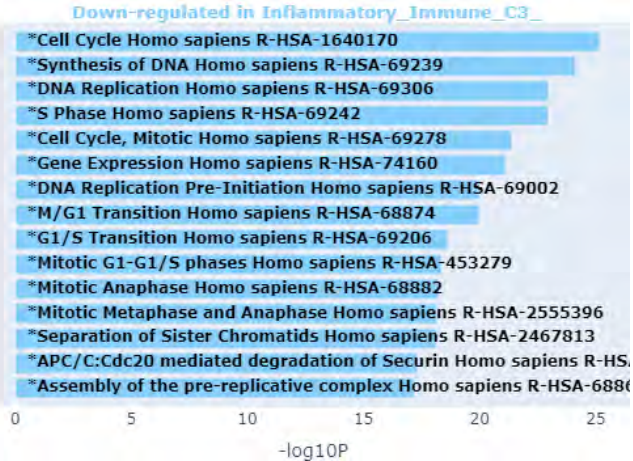
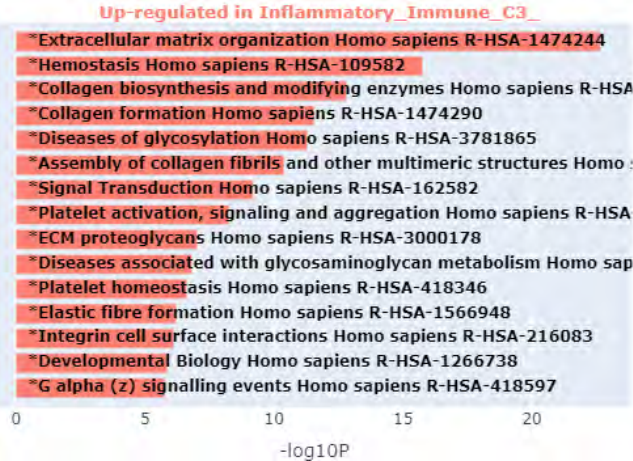


Gene symbol	logFC	AveExpr	t	P.Value
<i>nPRs</i>	1.510682	-2.04635	6.092938	3.45E-09
<i>PAQR8</i>	0.684746	1.709085	4.151653	4.32E-05
<i>PAQR5</i>	0.776807	1.385898	2.33205	2.04E-02
<i>CCM2</i>	-0.20795	5.975757	-2.57951	1.04E-02
<i>CCM3</i>	-0.1836	3.980731	-3.29313	1.11E-03

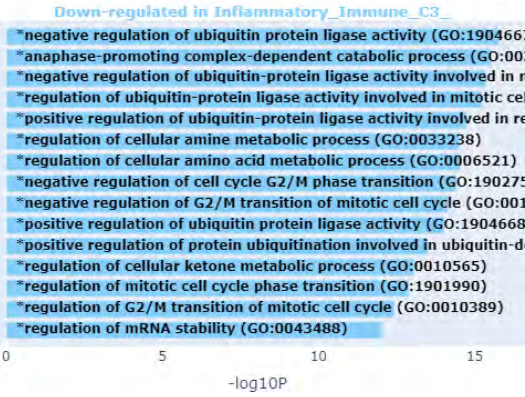
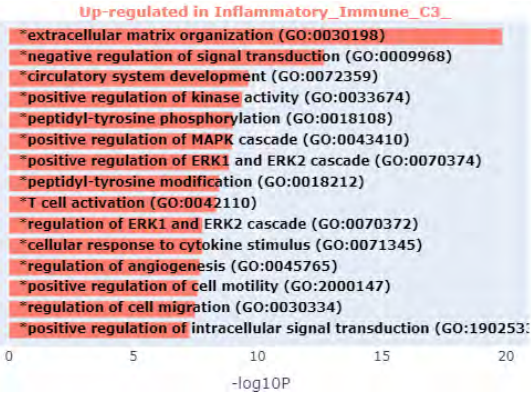
Molecular Functions



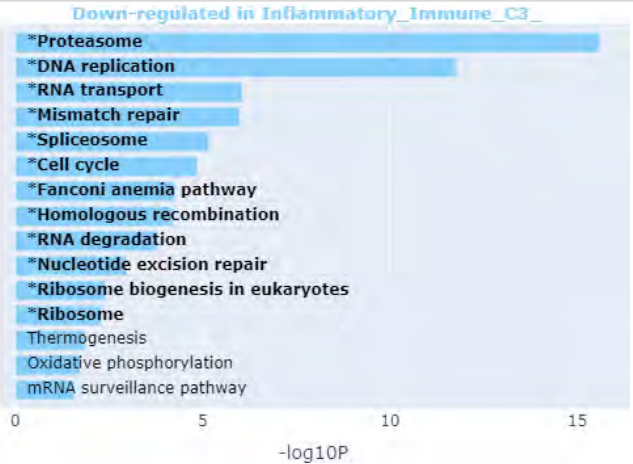
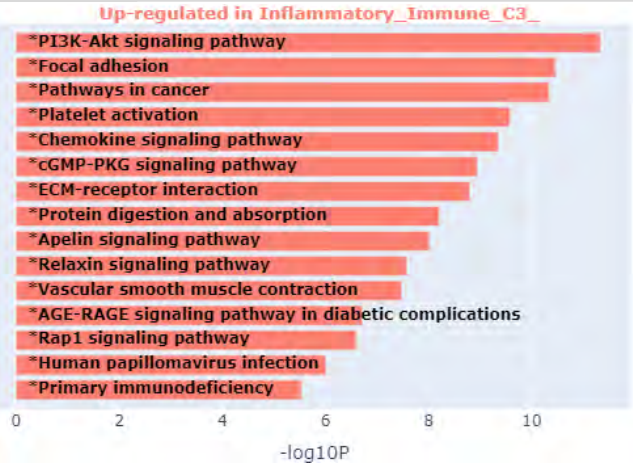
Reactome pathways



Biological Processes

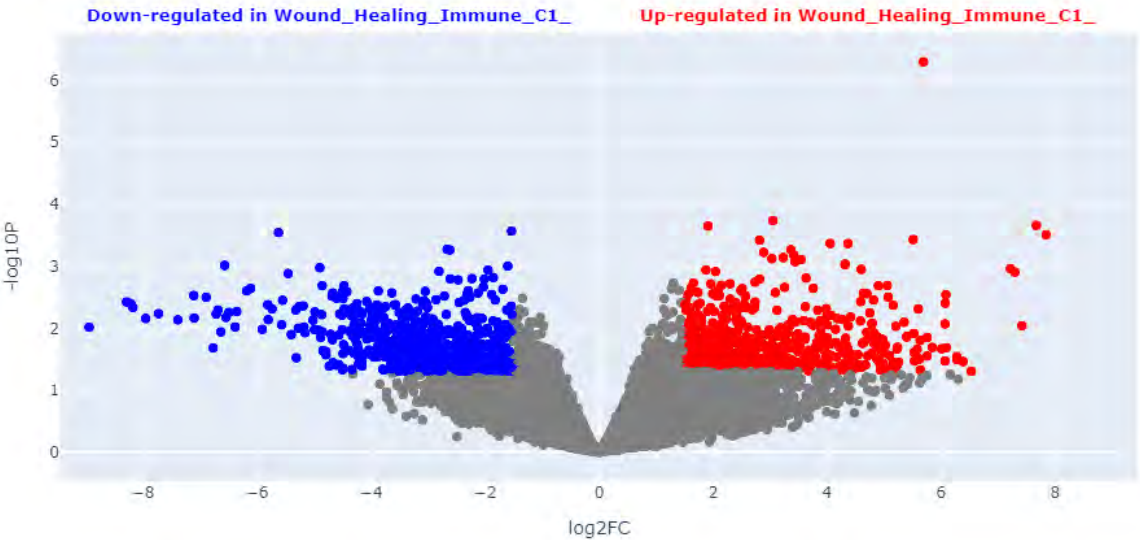


KEGG Pathways

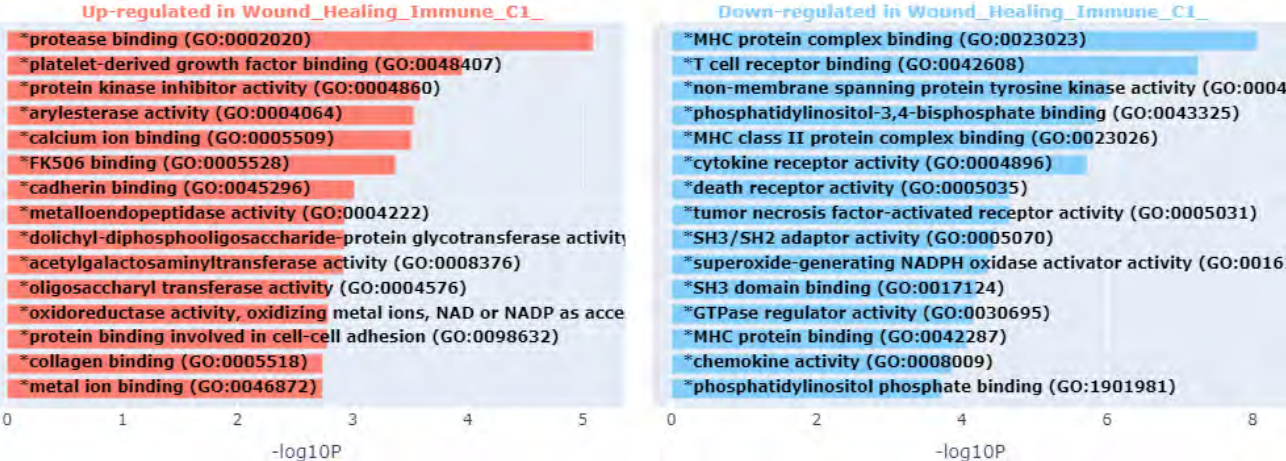


S2A

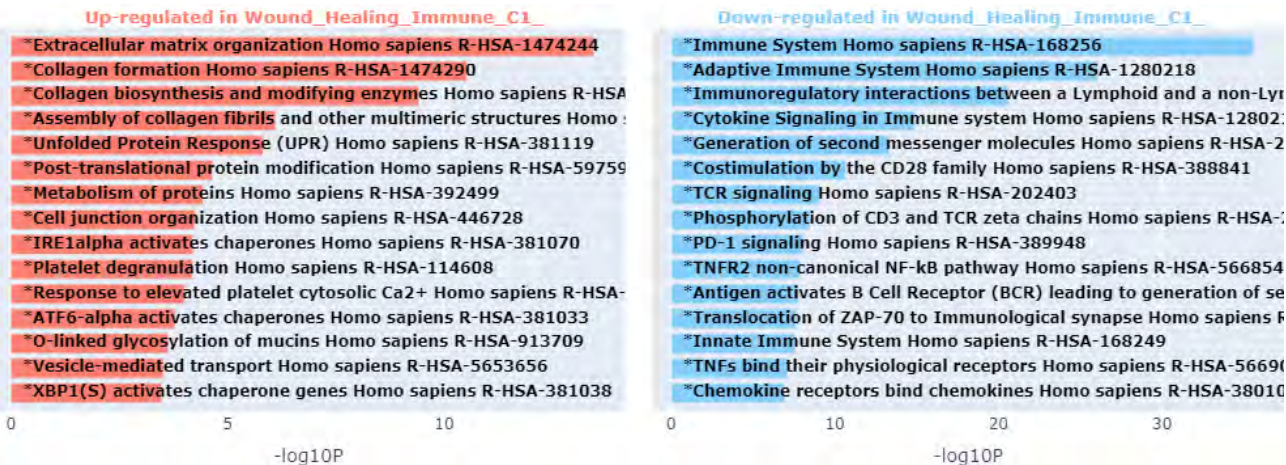
Immunological Subtypes	Sample size
Wound Healing (C1)	7
IFN-γ dominant (C2)	2



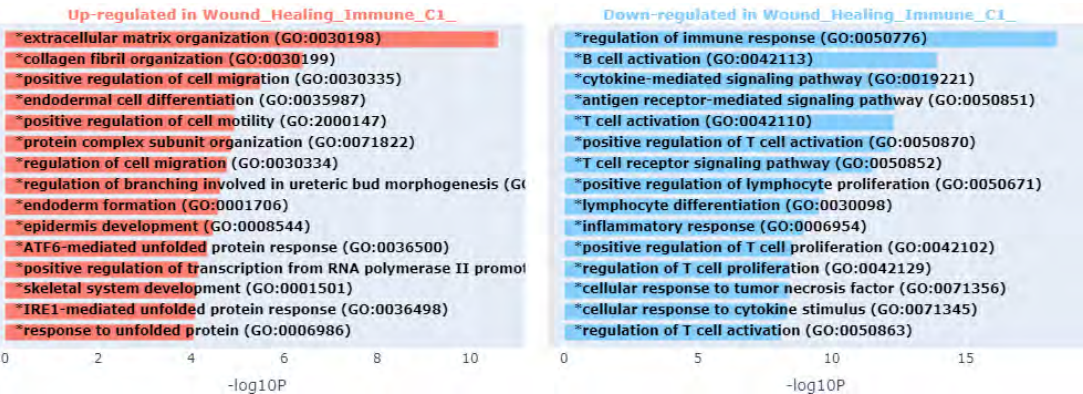
Molecular Functions



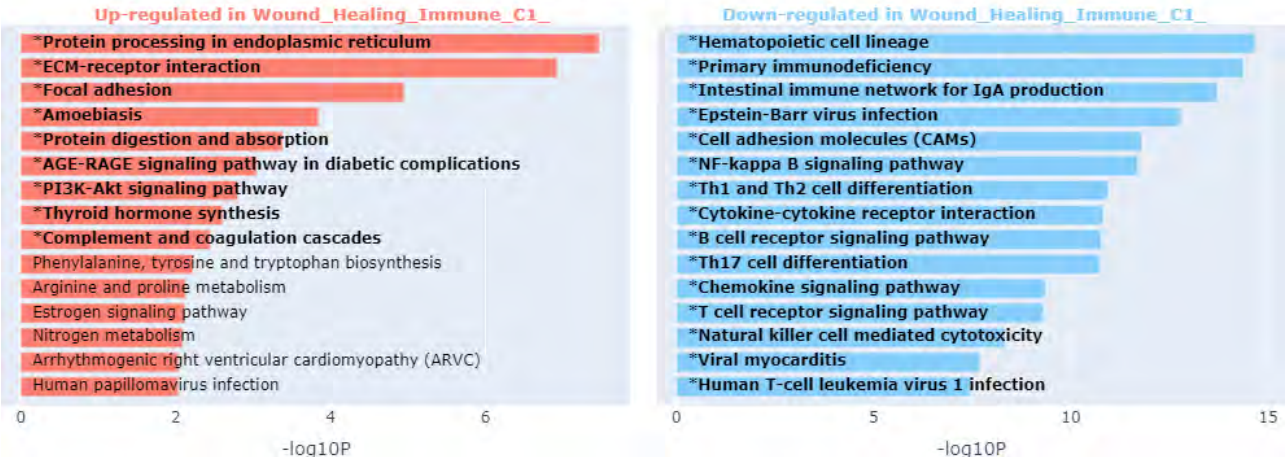
Reactome pathways



Biological Processes

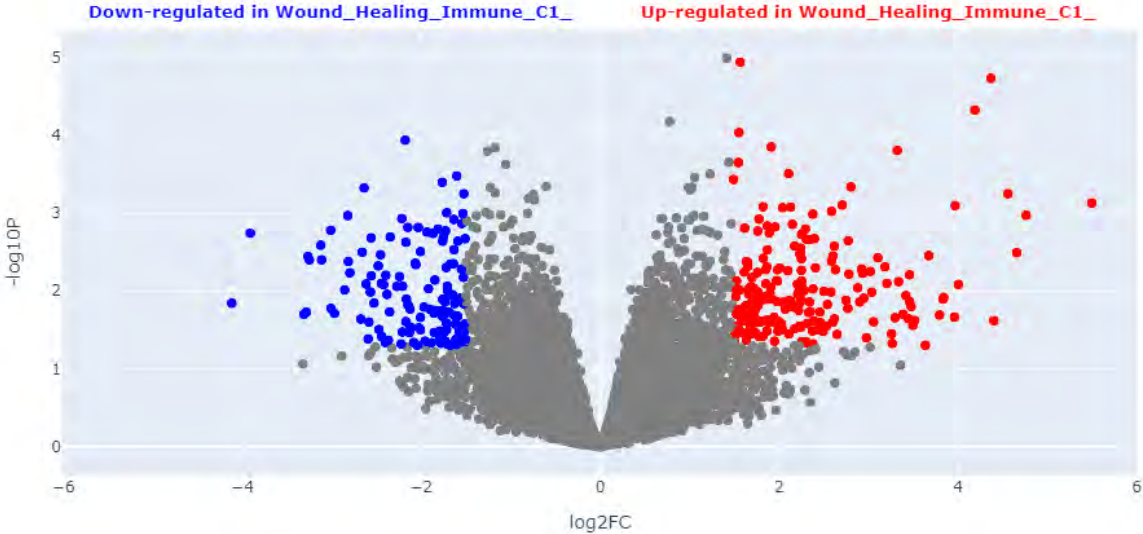


KEGG Pathways



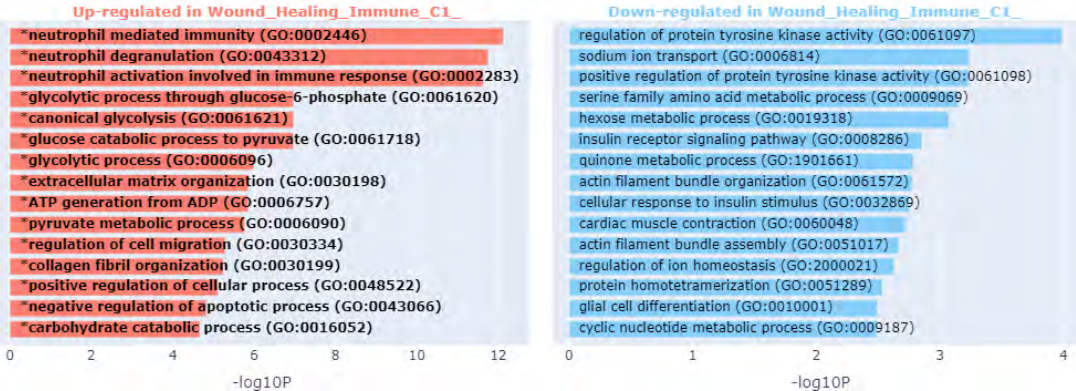
S2B

Immunological Subtypes	Sample size
Wound Healing (C1)	7
Inflammatory (C3)	17

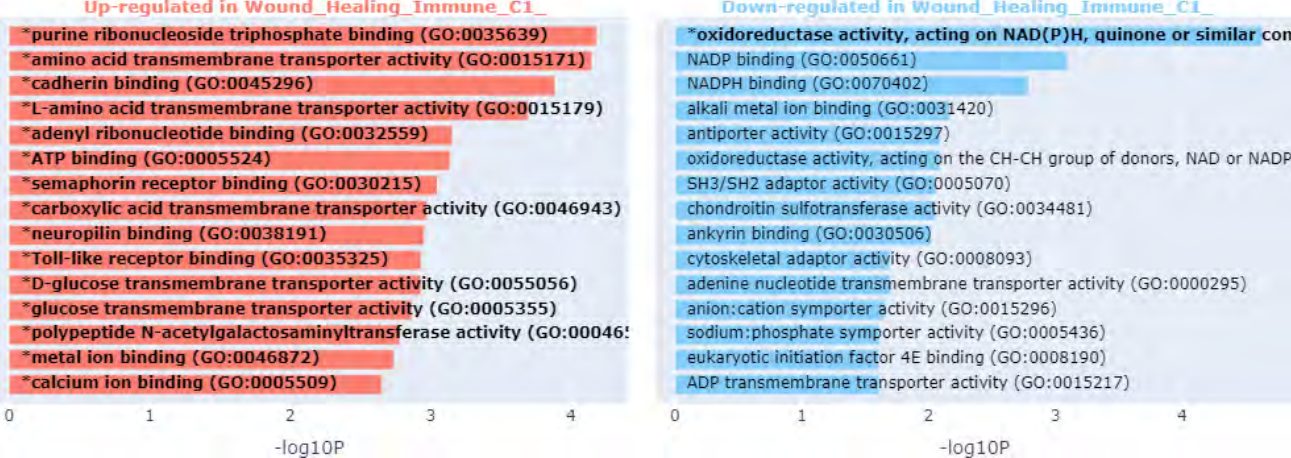


Gene symbol	logFC	AveExpr	t	P.Value
PAQR6	-1.3678	2.986596	-2.77353	1.02E-02

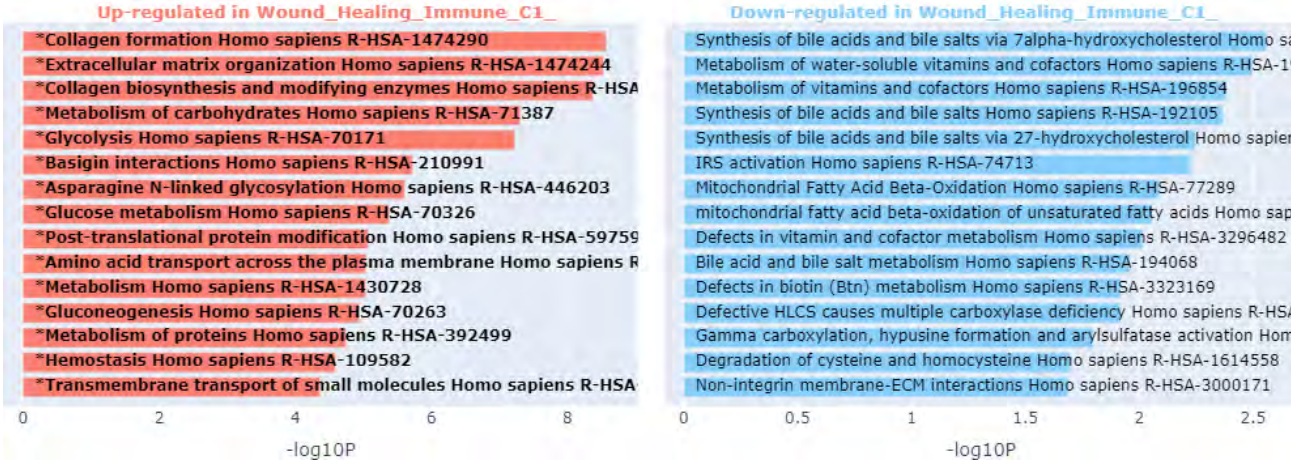
Biological Processes



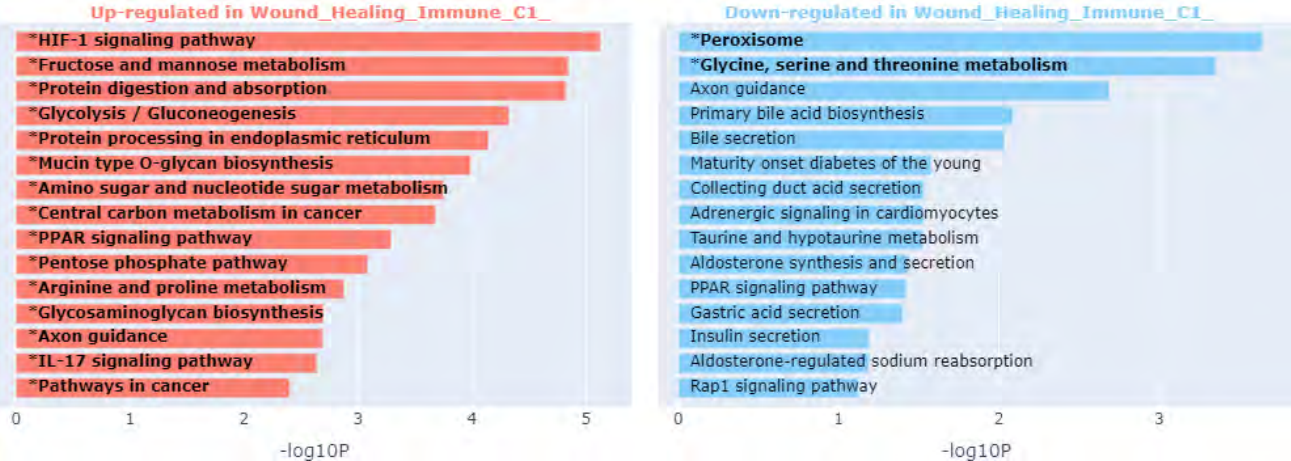
Molecular Functions



Reactome pathways

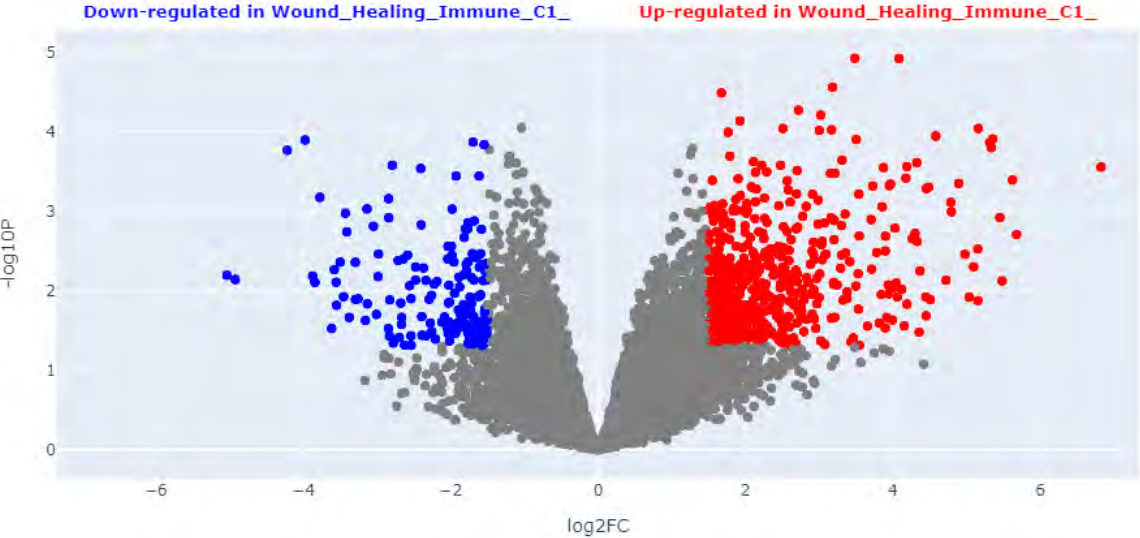


KEGG Pathways



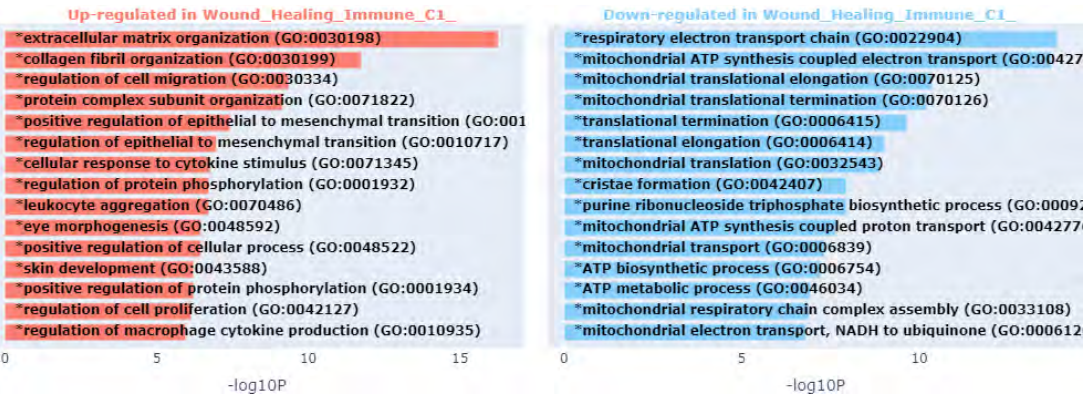
S2C

Immunological Subtypes	Sample size
Wound Healing (C1)	7
Lymphocyte depleted (C4)	8

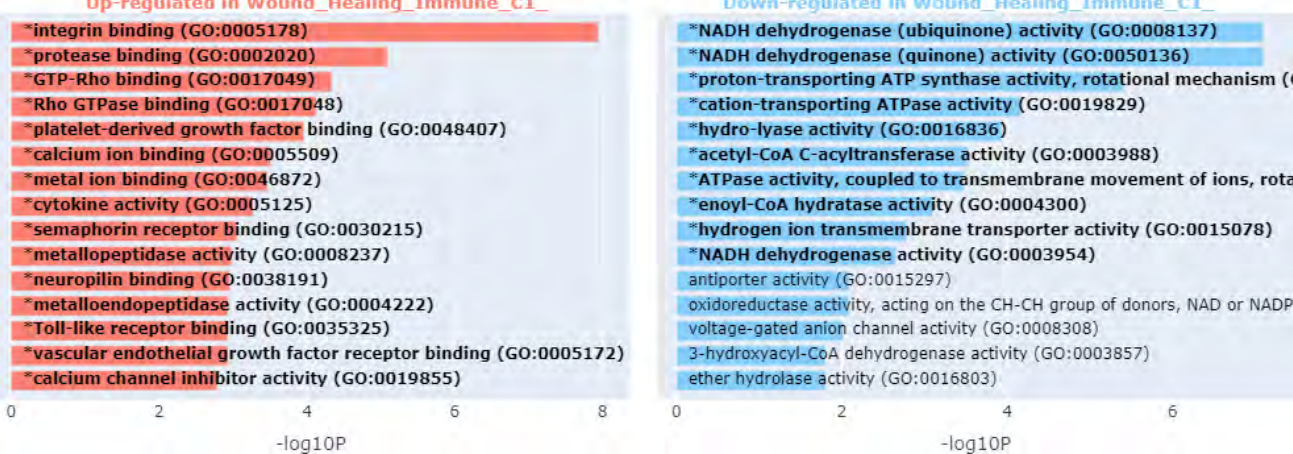


Gene symbol	logFC	AveExpr	t	P.Value
<i>PAQR6</i>	-1.47789	2.709641	-2.59804	1.90E-02

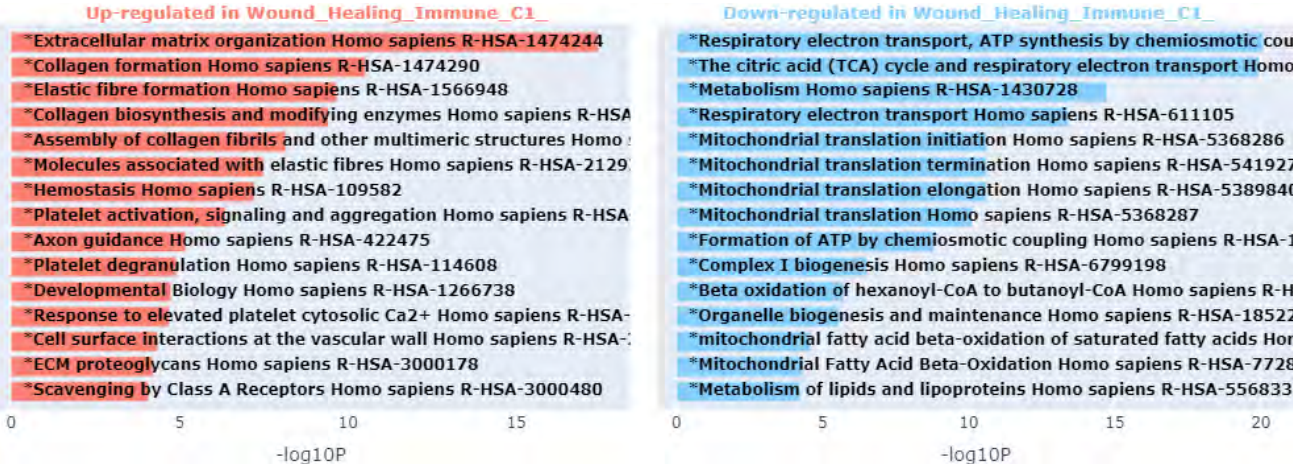
Biological Processes



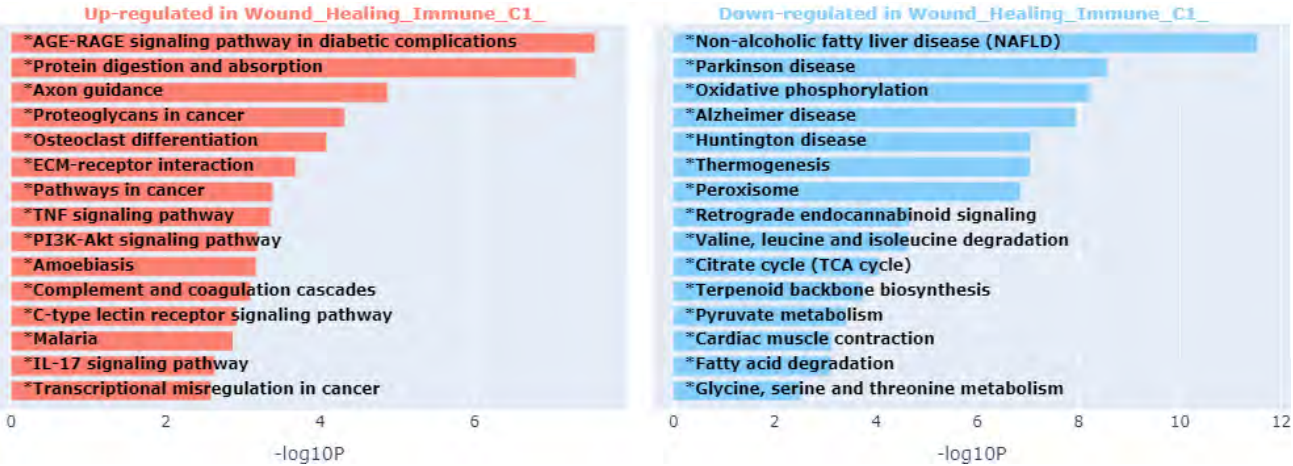
Molecular Functions



Reactome pathways

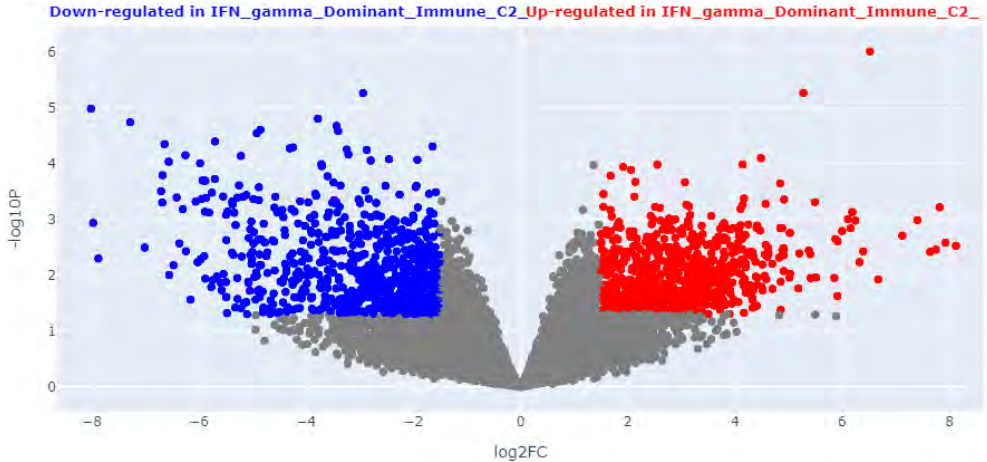


KEGG Pathways



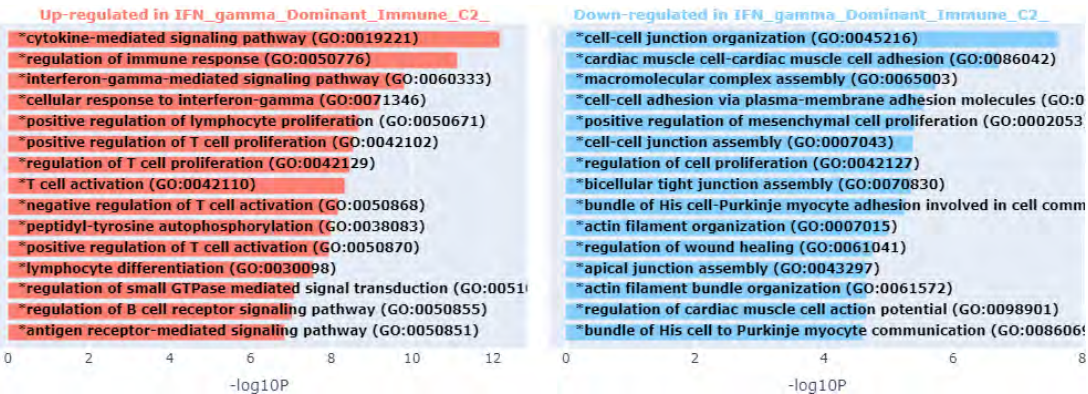
S2D

Immunological Subtypes	Sample size
IFN-γ dominant (C2)	2
Inflammatory (C3)	17

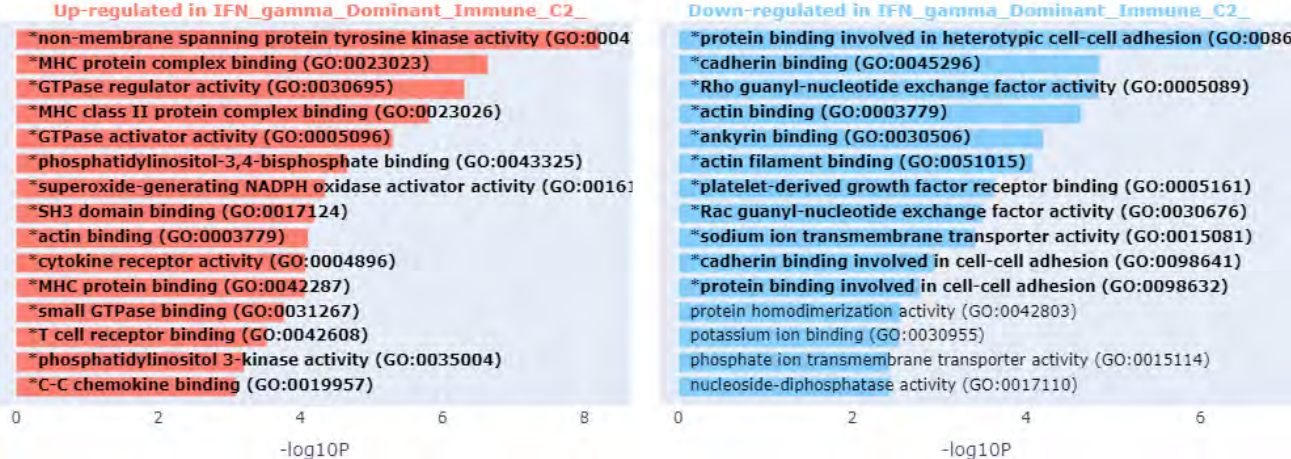


Gene symbol	logFC	AveExpr	t	P.Value
<i>CCM2</i>	1.031555	5.767136	2.317519	3.08E-02
<i>PAQR6</i>	-1.93451	3.191659	-2.32825	3.01E-02
<i>PGRMC1</i>	-1.08375	6.999647	-2.33136	2.99E-02
<i>PAQR5</i>	-3.83949	5.661036	-2.91322	8.40E-03

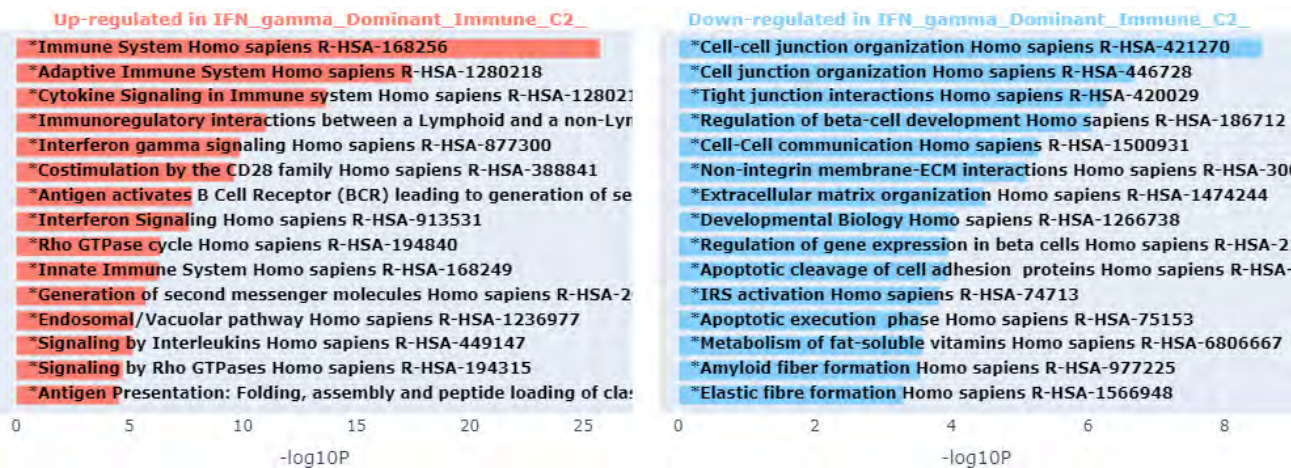
Biological Processes



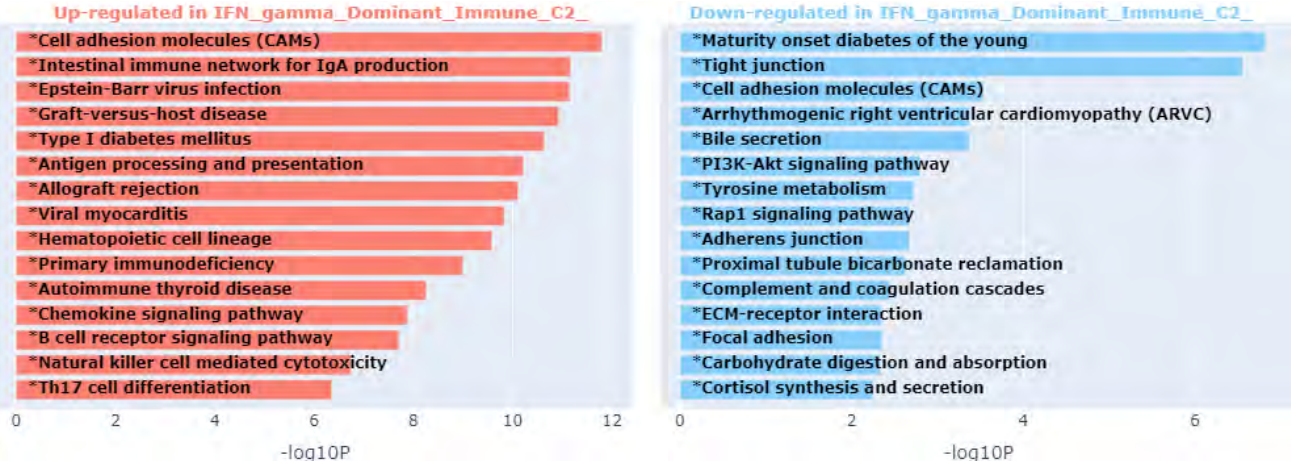
Molecular Functions



Reactome pathways



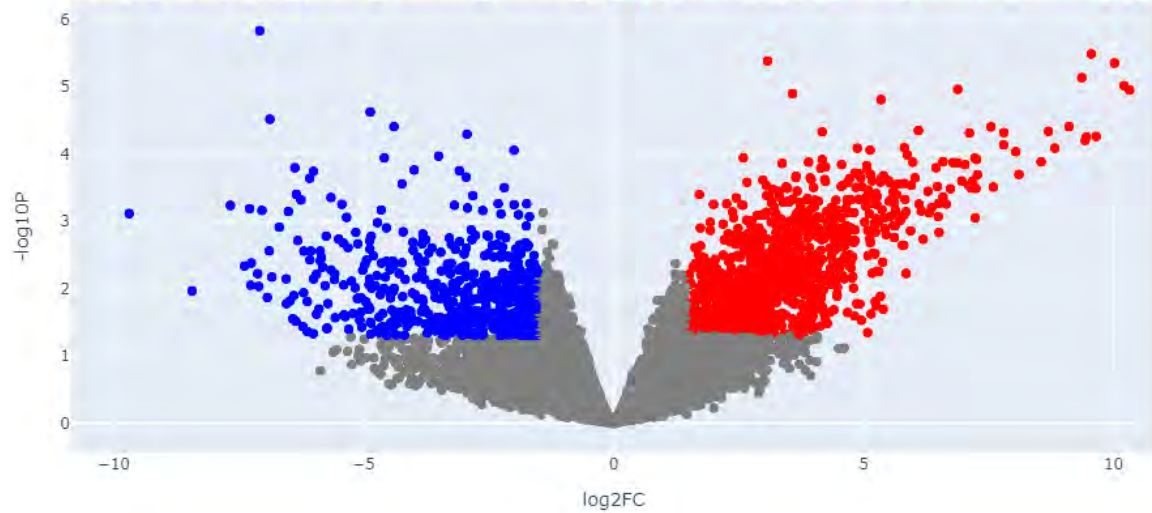
KEGG Pathways



S2E

Immunological Subtypes	Sample size
IFN-γ dominant (C2)	2
Lymphocyte depleted (C4)	8

Down-regulated in IFN_gamma_Dominant_Immune_C2 Up-regulated in IFN_gamma_Dominant_Immune_C2



Gene symbol	logFC	AveExpr	t	P.Value
<i>PGRMC1</i>	-1.49332	7.052058	-2.27509	4.27E-02
<i>PAQR6</i>	-2.11244	2.960783	-2.41878	3.30E-02

Biological Processes

Up-regulated in IFN_gamma_Dominant_Immune_C2

- *T cell activation (GO:0042110)
- *positive regulation of lymphocyte proliferation (GO:0050671)
- *regulation of immune response (GO:0050776)
- *antigen receptor-mediated signaling pathway (GO:0050851)
- *cytokine-mediated signaling pathway (GO:0019221)
- *positive regulation of T cell activation (GO:0050870)
- *regulation of T cell proliferation (GO:0042129)
- *inflammatory response (GO:0006954)
- *cellular response to cytokine stimulus (GO:0071345)
- *positive regulation of T cell proliferation (GO:0042102)
- *B cell activation (GO:0042113)
- *lymphocyte differentiation (GO:0030098)
- *T cell receptor signaling pathway (GO:0050852)
- *enzyme linked receptor protein signaling pathway (GO:0007167)
- *cellular response to Interferon-gamma (GO:0071346)

Down-regulated in IFN_gamma_Dominant_Immune_C2

- *ATF6-mediated unfolded protein response (GO:0036500)
- *response to endoplasmic reticulum stress (GO:0034976)
- *actin filament bundle assembly (GO:0051017)
- *regulation of actin filament polymerization (GO:0030833)
- *IRE1-mediated unfolded protein response (GO:0036498)
- *innate immune response in mucosa (GO:0002227)
- *cellular protein metabolic process (GO:0044267)
- *calcium-independent cell-cell adhesion via plasma membrane cell-cell contact (GO:0044267)
- *diterpenoid metabolic process (GO:0016101)
- *antibacterial humoral response (GO:0019731)
- *actin filament bundle organization (GO:0061572)
- *mucosal immune response (GO:0002385)
- *amyloid-beta metabolic process (GO:0050435)
- *negative regulation of epithelial to mesenchymal transition (GO:0007266)
- *Rho protein signal transduction (GO:0007266)

Molecular Functions

- *cytokine receptor activity (GO:0004896)
- *non-membrane spanning protein tyrosine kinase activity (GO:0004896)
- *MHC protein complex binding (GO:0023023)
- *MHC class II protein complex binding (GO:0023026)
- *chemokine receptor activity (GO:0004950)
- *MHC protein binding (GO:0042287)
- *phosphatidylinositol-3,4-bisphosphate binding (GO:0043325)
- *chemokine activity (GO:0008009)
- *MHC class II receptor activity (GO:0032395)
- *chemokine receptor binding (GO:0042379)
- *CCR chemokine receptor binding (GO:0048020)
- *protein tyrosine kinase activity (GO:0004713)
- *MHC class II protein binding (GO:0042289)
- *T cell receptor binding (GO:0042608)
- *GTPase regulator activity (GO:0030695)

- *peptide disulfide oxidoreductase activity (GO:0015037)
- *cadherin binding (GO:0045296)
- *oxidoreductase activity, acting on the CH-OH group of donors, NAD+ dependent (GO:0015037)
- *disulfide oxidoreductase activity (GO:0015036)
- *vascular endothelial growth factor receptor 2 binding (GO:0043184)
- *Rho guanyl-nucleotide exchange factor activity (GO:0005089)
- *cadherin binding involved in cell-cell adhesion (GO:0098641)
- *vascular endothelial growth factor receptor binding (GO:0005172)
- *intramolecular oxidoreductase activity, transposing S-S bonds (GO:0015037)
- *protein disulfide isomerase activity (GO:0003756)
- *sodium:phosphate symporter activity (GO:0005436)
- *protein binding involved in cell-cell adhesion (GO:0098632)
- *sodium ion transmembrane transporter activity (GO:0015081)
- *phosphate ion transmembrane transporter activity (GO:0015114)
- *anion:cation symporter activity (GO:0015296)

Reactome pathways

- *Immune System Homo sapiens R-HSA-168256
- *Adaptive Immune System Homo sapiens R-HSA-1280218
- *Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell Homo sapiens R-HSA-1280218
- *Translocation of ZAP-70 to Immunological synapse Homo sapiens R-HSA-1280218
- *Chemokine receptors bind chemokines Homo sapiens R-HSA-38010
- *Phosphorylation of CD3 and TCR zeta chains Homo sapiens R-HSA-38010
- *PD-1 signaling Homo sapiens R-HSA-389948
- *Costimulation by the CD28 family Homo sapiens R-HSA-388841
- *Cytokine Signaling in Immune system Homo sapiens R-HSA-1280218
- *Class A/1 (Rhodopsin-like receptors) Homo sapiens R-HSA-373076
- *TCR signaling Homo sapiens R-HSA-202403
- *Innate Immune System Homo sapiens R-HSA-168249
- *GPVI-mediated activation cascade Homo sapiens R-HSA-114604
- *G alpha (i) signalling events Homo sapiens R-HSA-418594

- *Unfolded Protein Response (UPR) Homo sapiens R-HSA-381119
- *Metabolism of proteins Homo sapiens R-HSA-392499
- *Amyloid fiber formation Homo sapiens R-HSA-977225
- *Metabolism Homo sapiens R-HSA-1430728
- *ATF6-alpha activates chaperones Homo sapiens R-HSA-381033
- *Tight junction interactions Homo sapiens R-HSA-420029
- *RNA Polymerase I Promoter Opening Homo sapiens R-HSA-73728
- *IRE1alpha activates chaperones Homo sapiens R-HSA-381070
- *Packaging Of Telomere Ends Homo sapiens R-HSA-171306
- *DNA methylation Homo sapiens R-HSA-5334118
- *DNA Damage/Telomere Stress Induced Senescence Homo sapiens R-HSA-421270
- *Activated PKN1 stimulates transcription of AR (androgen receptor) Homo sapiens R-HSA-421270
- *Cell-cell junction organization Homo sapiens R-HSA-421270
- *SIRT1 negatively regulates rRNA Expression Homo sapiens R-HSA-421270
- *XBP1(S) activates chaperone genes Homo sapiens R-HSA-381038

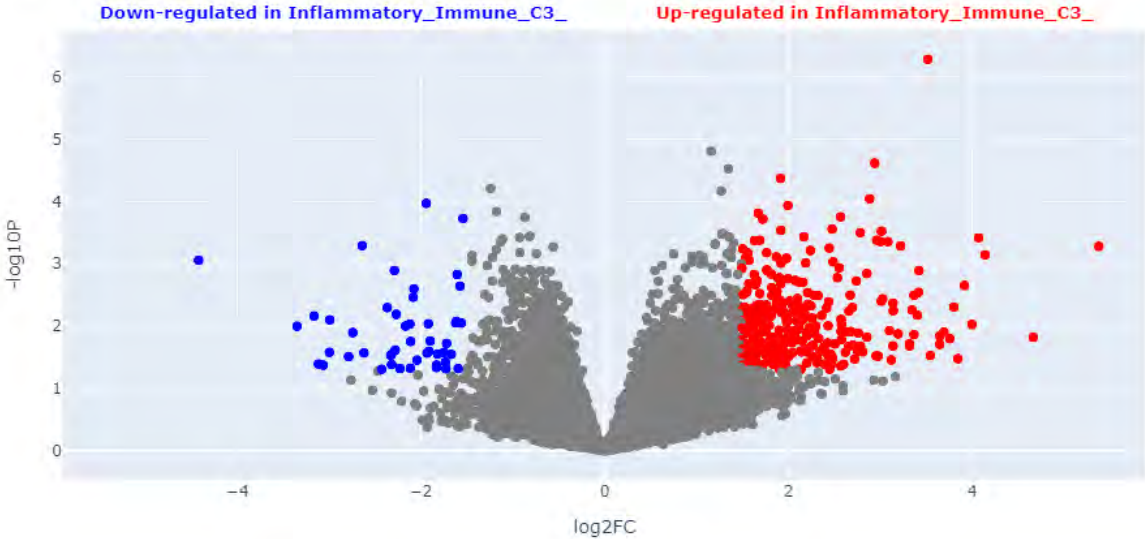
KEGG Pathways

- *Hematopoietic cell lineage
- *Th17 cell differentiation
- *Th1 and Th2 cell differentiation
- *Chemokine signaling pathway
- *Cell adhesion molecules (CAMs)
- *Cytokine-cytokine receptor interaction
- *Primary immunodeficiency
- *Intestinal immune network for IgA production
- *Inflammatory bowel disease (IBD)
- *Leishmaniasis
- *Allograft rejection
- *T cell receptor signaling pathway
- *Graft-versus-host disease
- *Natural killer cell mediated cytotoxicity
- *Tuberculosis

- *Protein processing in endoplasmic reticulum
- *Maturity onset diabetes of the young
- *Thyroid hormone synthesis
- *Valine, leucine and isoleucine degradation
- *Tight junction
- *Pyruvate metabolism
- *Fatty acid degradation
- *Insulin resistance
- *Systemic lupus erythematosus
- *Arginine and proline metabolism
- *Other glycan degradation
- *Bacterial invasion of epithelial cells
- *Cell adhesion molecules (CAMs)
- *Tyrosine metabolism
- *Non-alcoholic fatty liver disease (NAFLD)

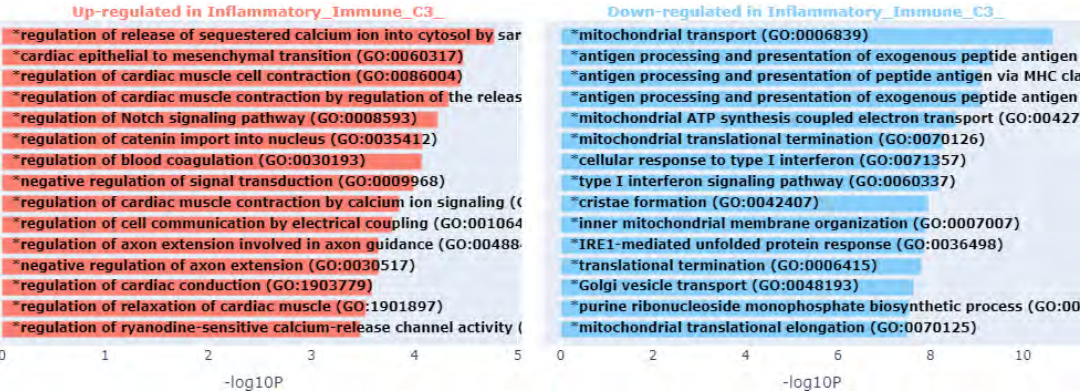
S2F

Immunological Subtypes	Sample size
Inflammatory (C3)	17
Lymphocyte depleted (C4)	8

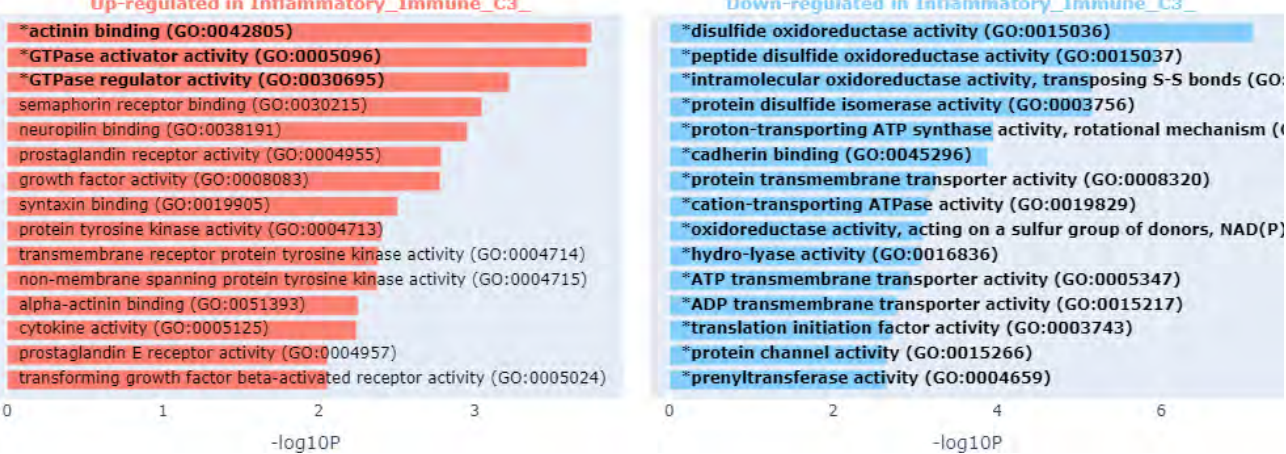


Gene symbol	logFC	AveExpr	t	P.Value
CCM3	-0.56277	4.842558	-2.64929	1.34E-02

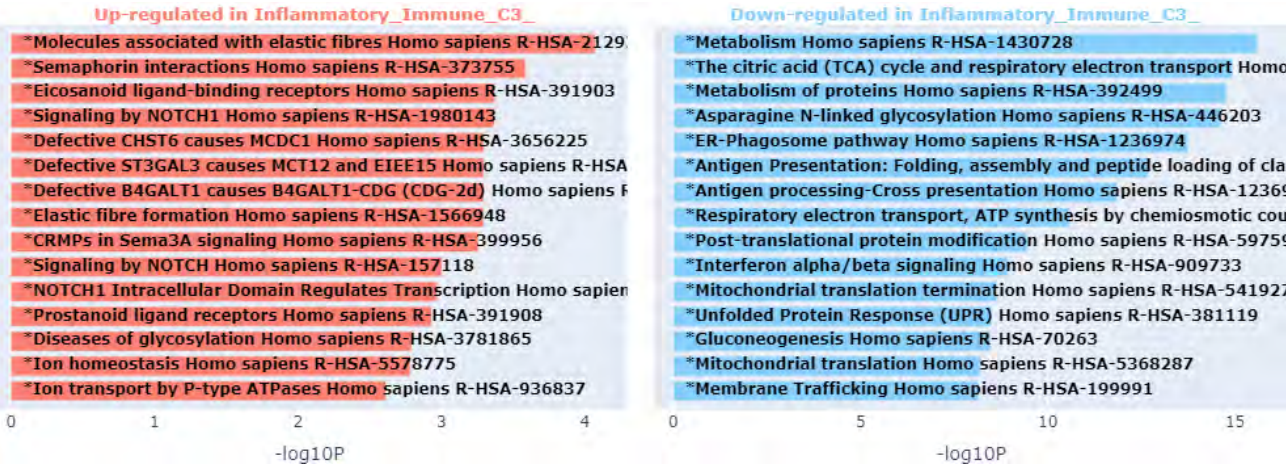
Biological Processes



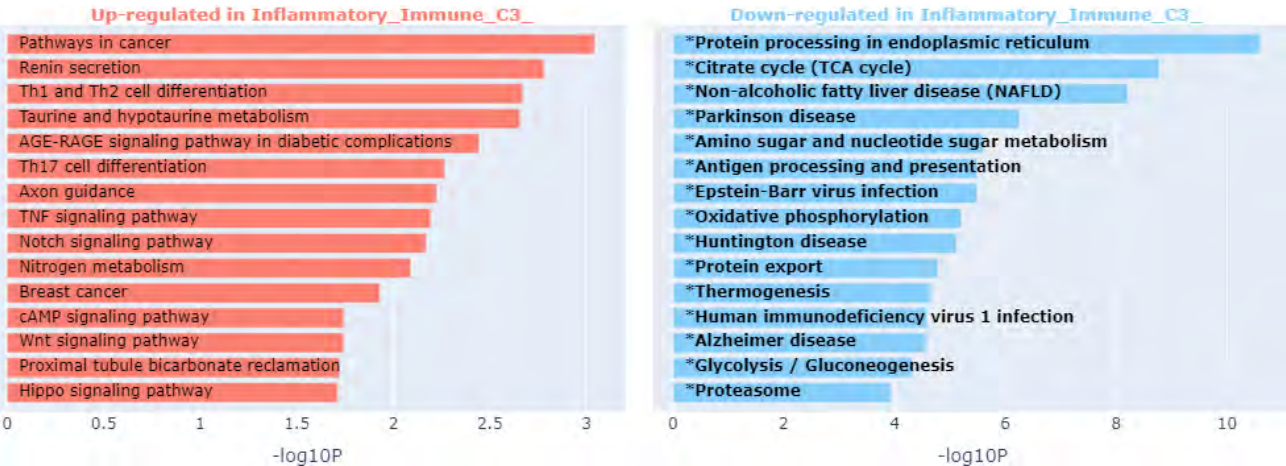
Molecular Functions



Reactome pathways

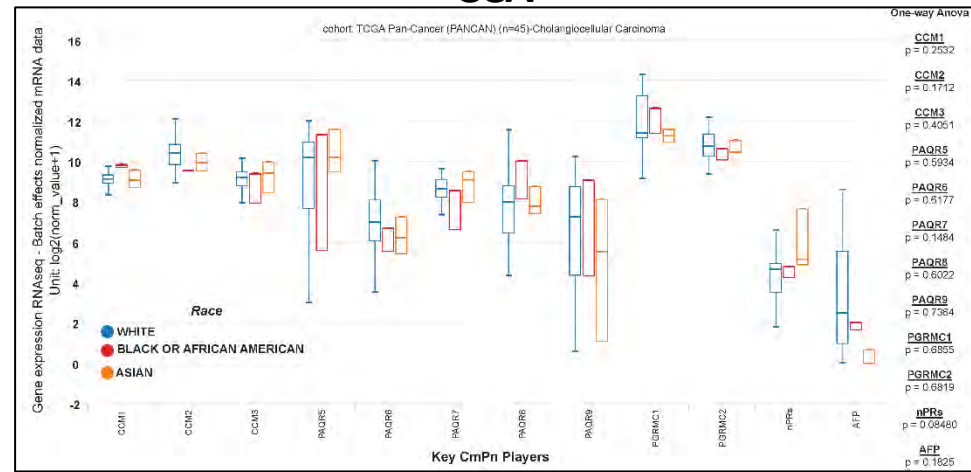


KEGG Pathways

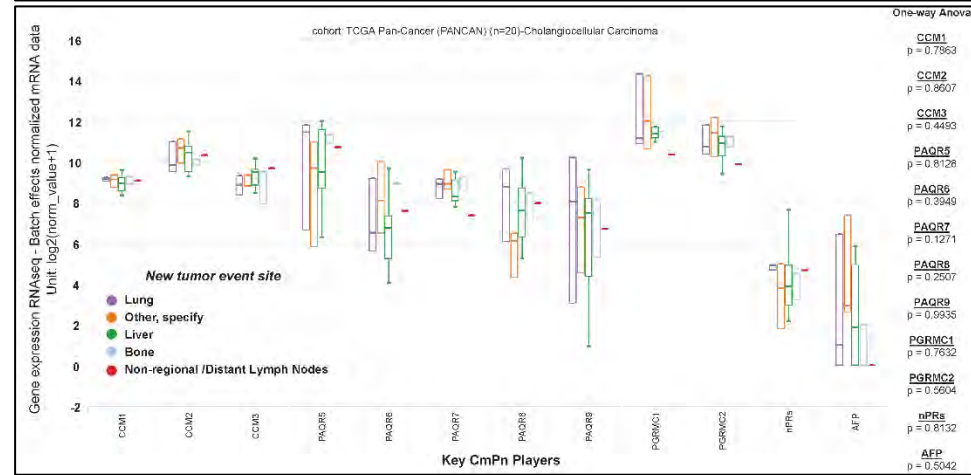


S3A.

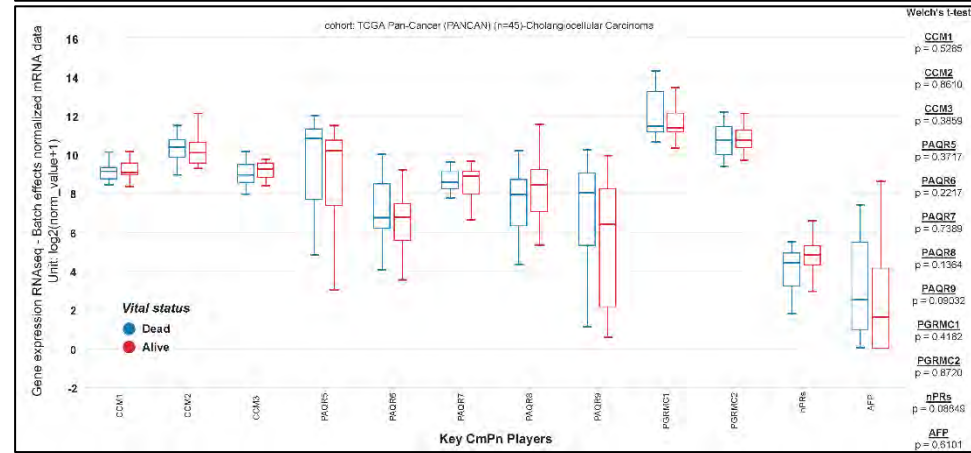
CCA



B.

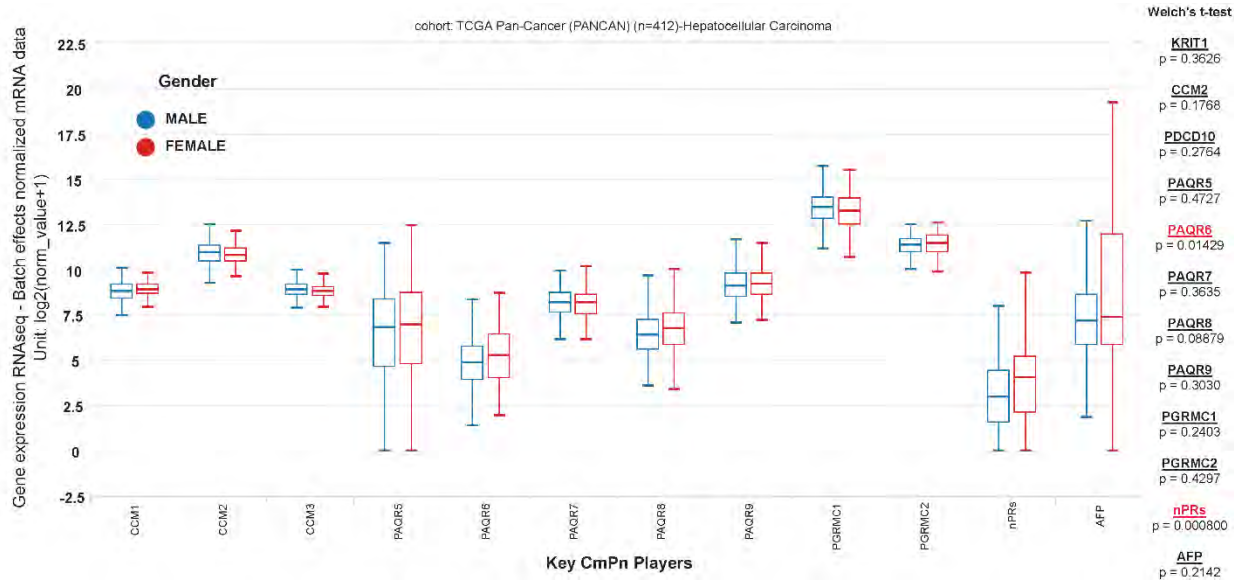


C.



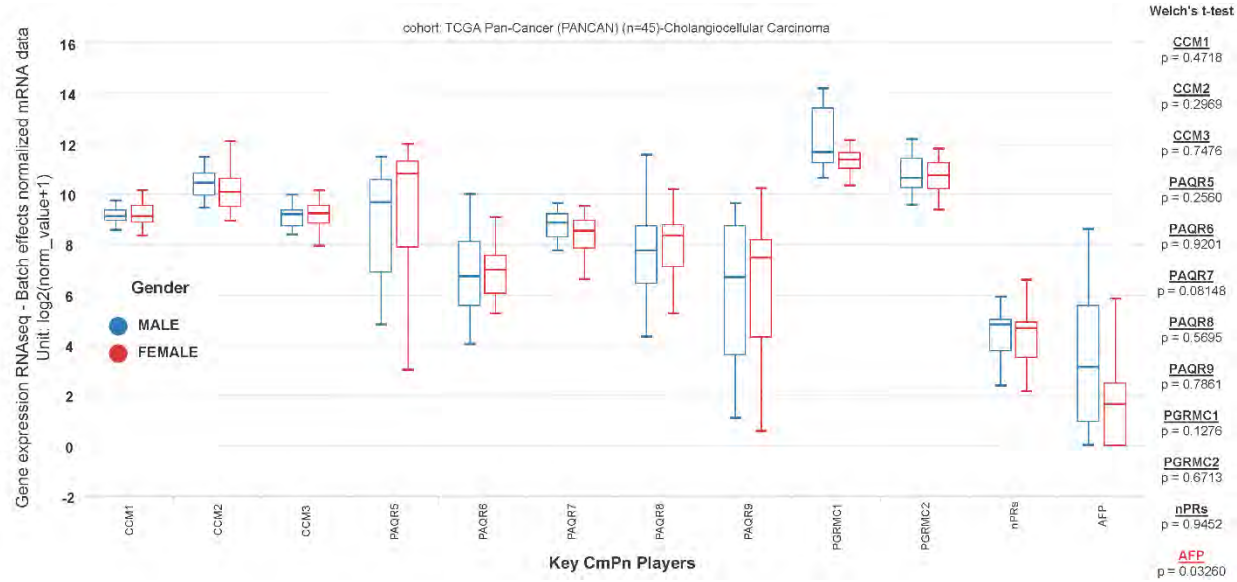
D-1.

HCC

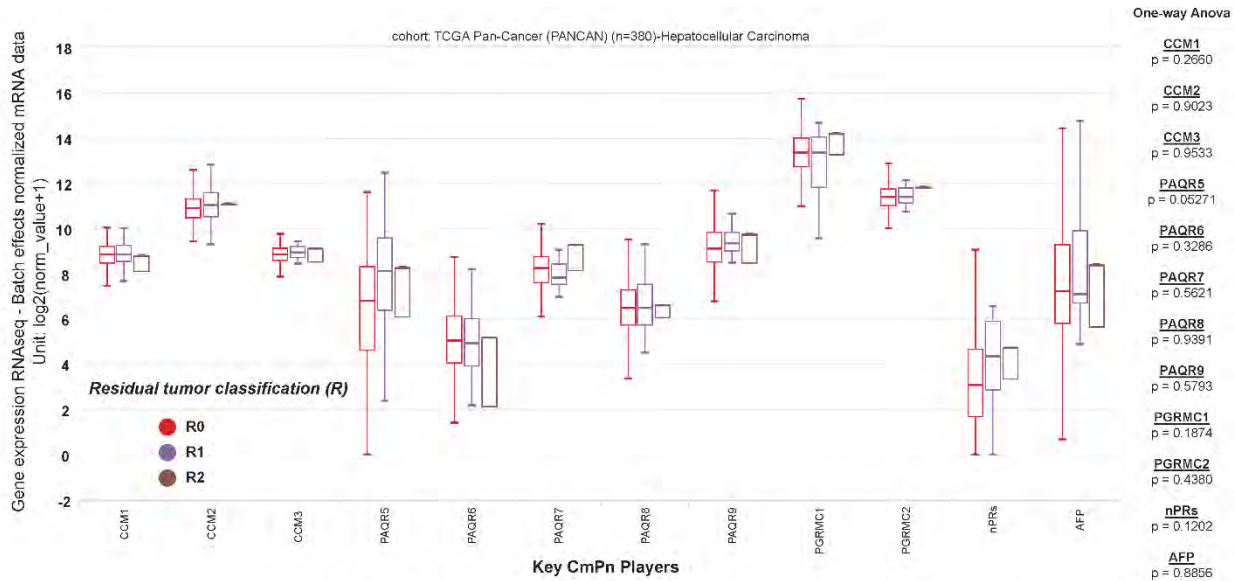


D-2.

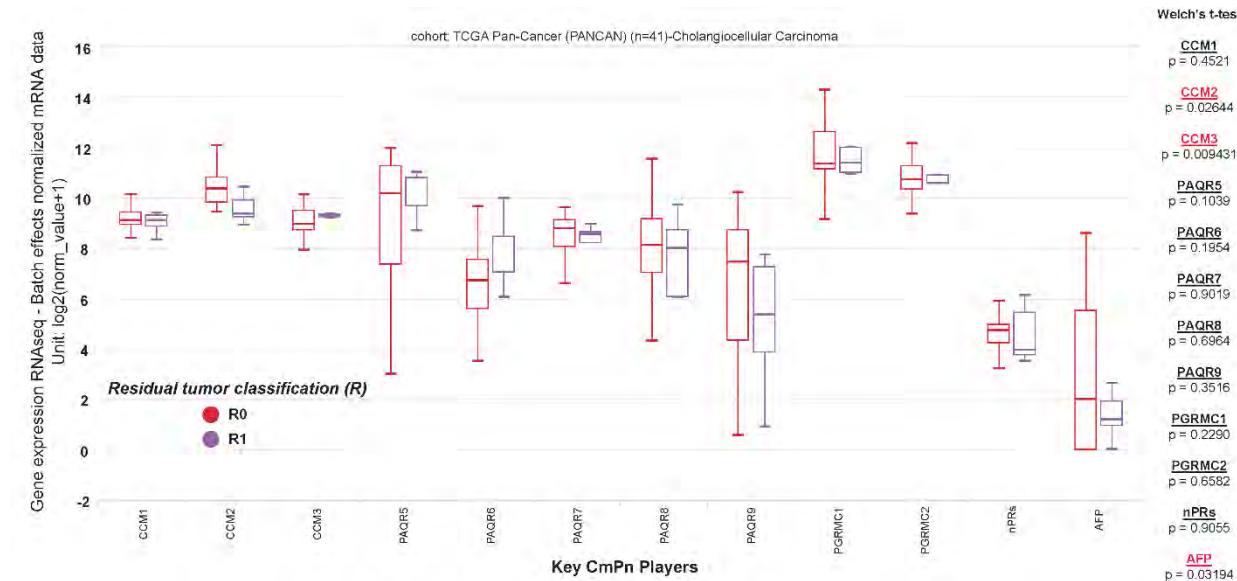
CCA

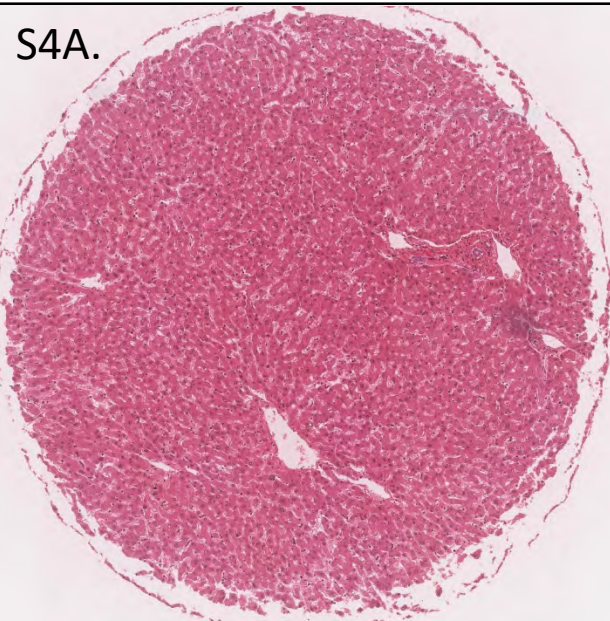


E-1.

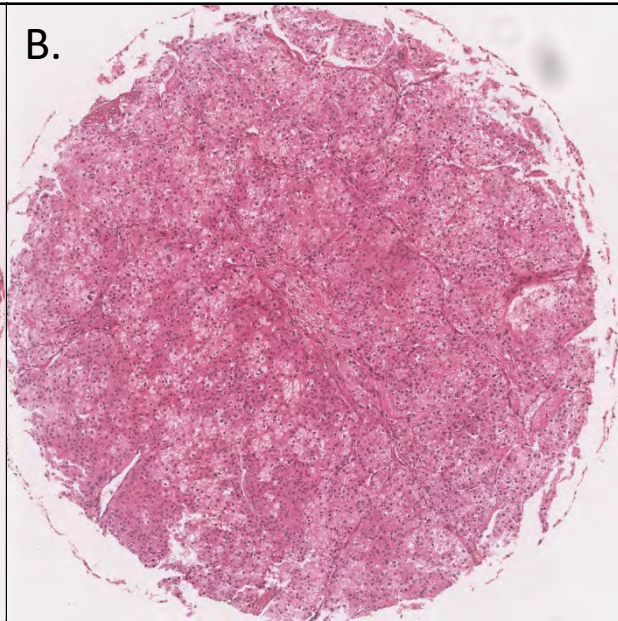


E-2.

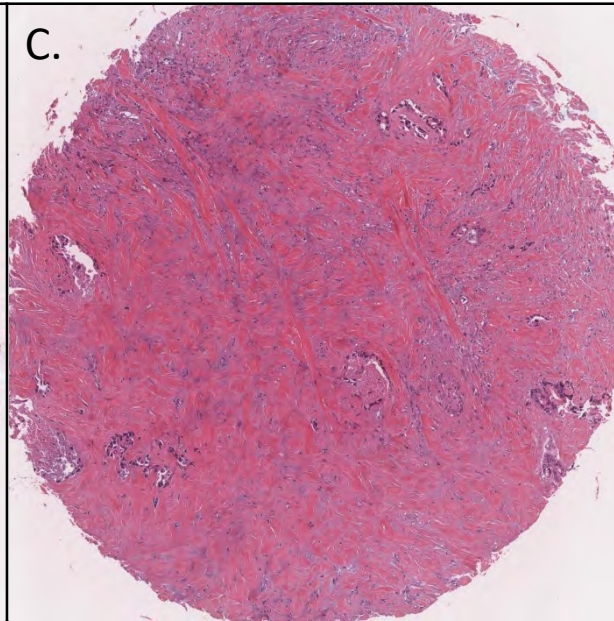




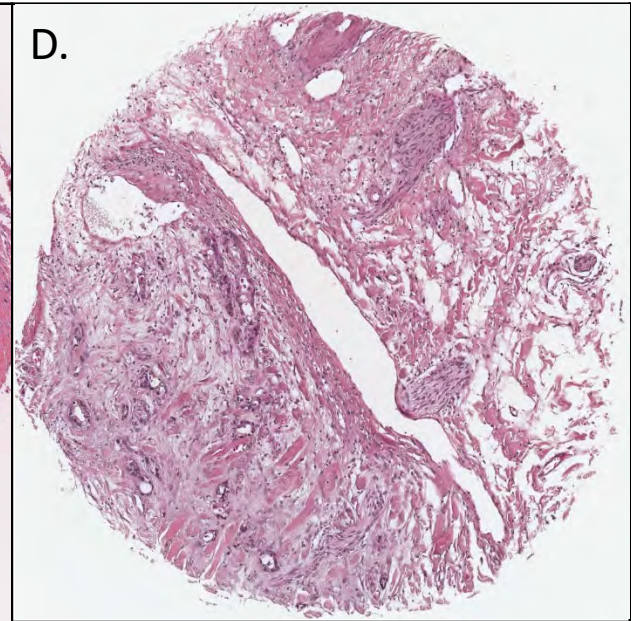
Normal liver tissue (NORM)



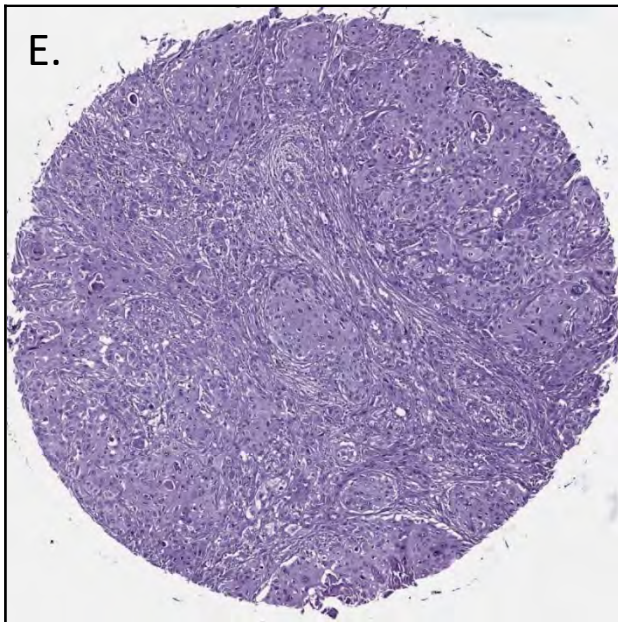
Hepatocellular Carcinoma (HCC)



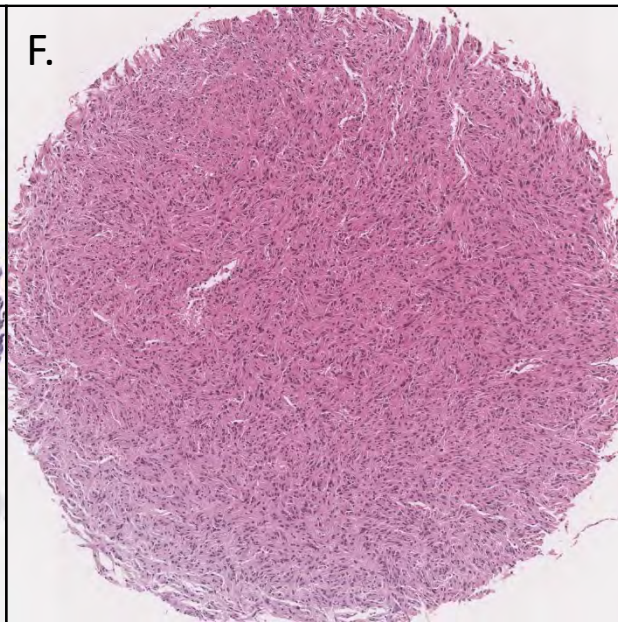
Cholangiocarcinoma (CCA)



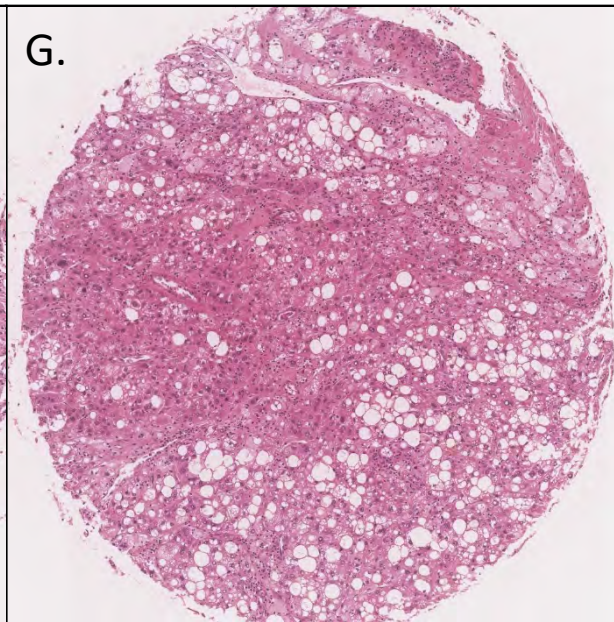
Hepatocholangiocarcinoma (cHCC-CCA)



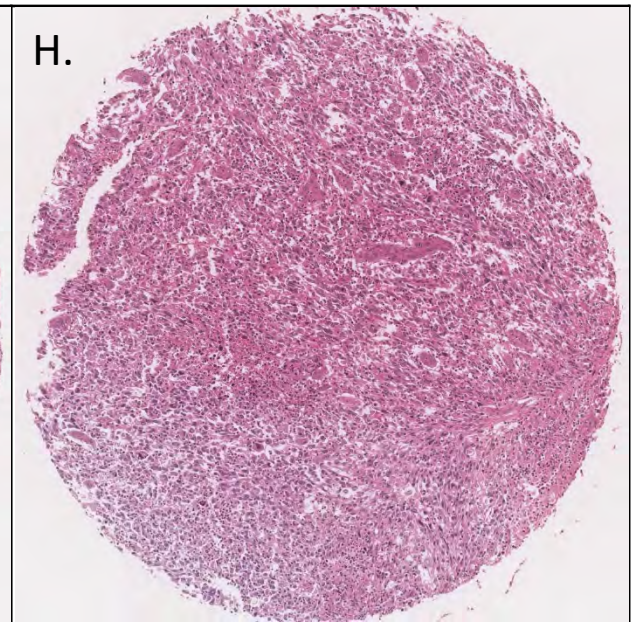
Adenosquamous Carcinoma (ASC)



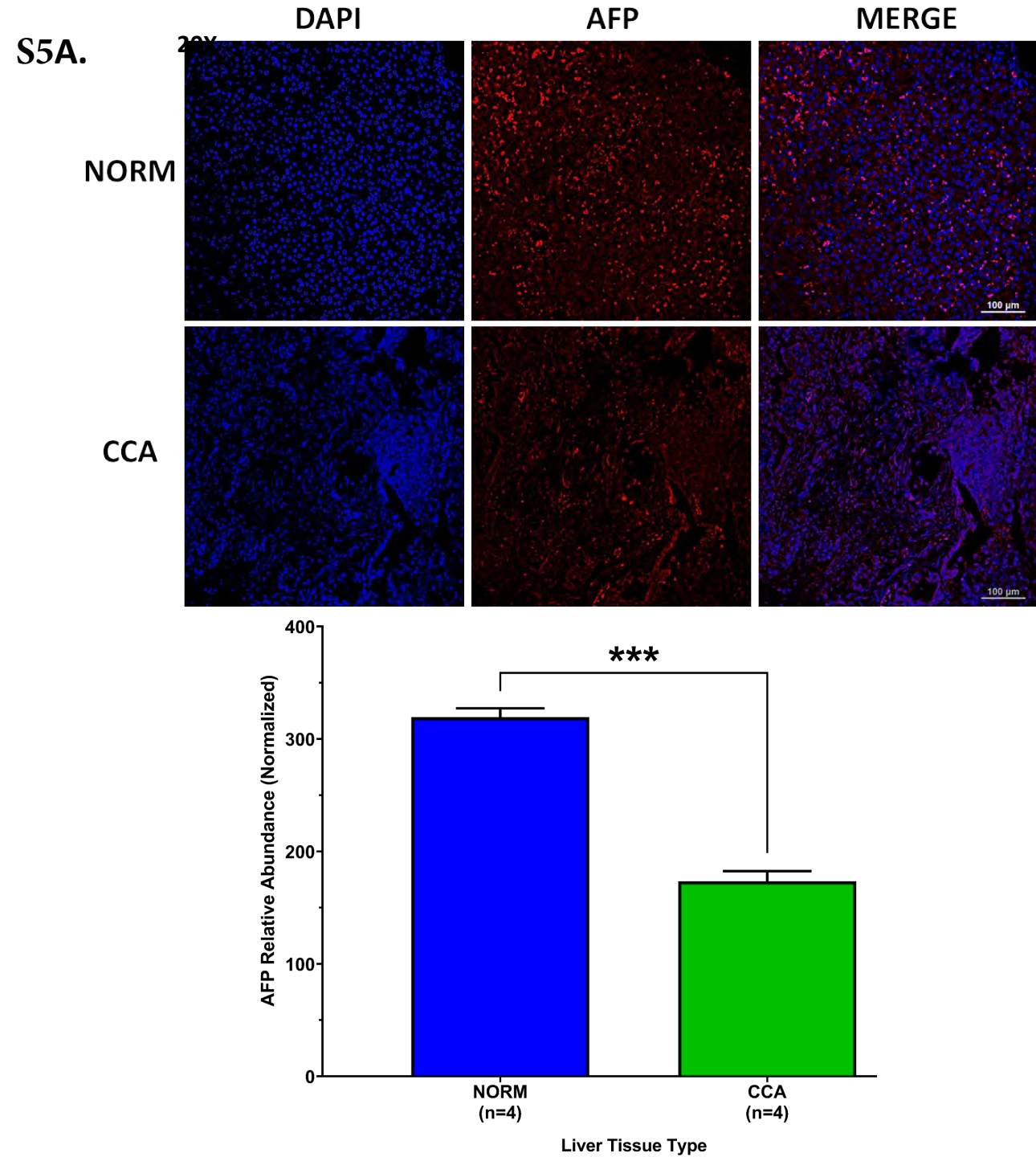
Undifferentiated Pleomorphic Sarcoma (UPS)



Clear Cell Carcinoma (CCC)



Hepatic Angiosarcoma (HAS)



B.

nPRs Quantification Summary

